Economics Department

Likelihood Analysis of Seasonal Cointegration

SØREN JOHANSEN
and
ERNST SCHAUMBURG

ECO No. 97/16

EUI WORKING PAPERS

EUROPEAN UNIVERSITY INSTITUTE

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 97/16

Likelihood Analysis of Seasonal Cointegration

SØREN JOHANSEN
and
ERNST SCHAUMBURG

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in May 1997
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

Likelihood analysis of seasonal cointegration

S. Johansen
European University Institute
and
E. Schaumburg
Princeton University

March 14, 1997

Abstract

The vector autoregressive model for seasonal cointegration is analysed. The general error correction model is discussed and conditions are found under which the process is integrated of order 1 at seasonal frequency and exhibits cointegration.

Under these conditions a representation theorem for the solution is given expressed in terms of seasonal random walks. Finally the asymptotic properties of the likelihood ratio test for cointegrating rank is given, and it is shown that the estimated cointegrating vectors are asymptotically mixed Gaussian. The results resemble the result for cointegration at zero frequency but expressed in terms of a complex Brownian motion. Tables are provided for asymptotic inference under various assumptions on the deterministic terms.

Keywords: Autoregressive process; Granger's theorem; Error correction model; Complex Brownian motion.

1 Introduction

This paper contains a systematic treatment of the statistical analysis of seasonal cointegration in the vector autoregressive model. The theory started with the paper by Hylleberg, Engle, Granger and Yoo (1990) which gave the main results on the representation and the univariate tests for cointegration at the various complex frequencies.

The paper on maximum likelihood inference by Lee (1992) set the stage for the analysis of multivariate systems. Unfortunately it does not treat all aspects of asymptotic inference, and the test for cointegration rank is only partially correct. The two papers by Gregoir ($1993a, b$) deal with a very general situation of unit roots allowing for processes to be integrated of order greater than 1, but do not treat likelihood inference.

The purpose of this paper is therefore to improve on the previous analysis and discuss maximum likelihood estimation, calculation of test statistics, and derivation of asymptotic distributions in the context of the vector autoregressive model. In the process of doing so it is natural to give the mathematical theory of the Granger representation, the error correction model, the role of the constant term and seasonal dummies, as well as the asymptotic distributions involved. The basic new trick is the introduction of the complex Brownian motion, which makes many of the calculations more natural and greatly simplifies the formulae for the limit distributions.

The model we are considering is thus the autoregressive model defined for a $p-$ dimensional process X_t by the equations

$$X_t = \sum_{j=1}^{l} \Pi_j X_{t-j} + \Phi D_t + \varepsilon_t, \tag{1}$$

where we assume that the initial values $X_0, \ldots, X_{-l+1}$ are fixed, and that the errors are i.i.d. with mean zero and variance Ω. In the derivations of estimators and test statistics we also assume that ε_t are $N_p(0, \Omega)$, but in the derivation of the asymptotic results this assumption is not needed. The deterministic terms D_t can contain a constant, a linear term or seasonal dummies. Various models defined by restrictions on the deterministic terms will be considered. The properties of the process generated by the equations (1) are as usual expressed in terms of the characteristic polynomial

$$A(z) = I - \sum_{j=1}^{l} \Pi_j z^j,$$

with determinant $|A(z)|$.

The paper is organized as follows, in Section 2 we discuss the error correction model for seasonal cointegration of processes that are integrated of order 1 at seasonal frequency. This model is subsequently solved in the form of a Granger representation theorem, applying a general theorem about inversion of matrix polynomials. It is shown how this can be used to analyze the role of constant, linear term, and seasonal dummies, see Kunst and Franses (1994). In Section 3 we discuss the Gaussian likelihood analysis and the calculation of the maximum likelihood estimator in the model with unrestricted deterministic terms, as well as in some models defined by restrictions of deterministic terms. In Section 4 some technical asymptotic results on the behaviour of the processes and the product moments are given. Section 5 contains the asymptotic results for the maximum likelihood estimator of the cointegrating vectors, and the asymptotic distribution of the likelihood ratio test for cointegration rank at seasonal frequency are given.

In Appendix A a brief description of the matrix representation of complex matrices is given along with proofs of the technical results in Section 4. Finally Appendix B has the tables of the limit distribution of the likelihood ratio test for cointegrating rank for various models defined by restrictions on the deterministic terms.

2 The representation theorem and the error correction model

In this section we first give the necessary analytic results from the theory of polynomials $A(z)$ with values in the set of $p \times p$ matrices. In Theorem 1 we discuss Laplace's expansion for a polynomial around many points and show in Corollary 2 how this contains the formulation of an error correction model. The basic result, however, is Theorem 3 which is a necessary and sufficient condition for the inverse matrix polynomial to have poles of order 1. In Theorem 4 we give the interpretation of this result as a representation of the solution of the autoregressive equations allowing for integrated processes and cointegration at seasonal frequency, corresponding to Granger's theorem for $I(1)$ processes. We apply the Granger representation theorem to discuss the role of constant, linear term, and in particular seasonal dummies. We conclude this section with some examples of models for annual, semi-annual and quarterly data.

We are concerned with the roots of the equation $|A(z)| = 0$ in particular the unit roots, for which $|z| = 1$. For a complex number $z_1 = e^{i\theta}$, the complex conjugate is the same as the inverse, $\bar{z}_1 = z_1^{-1} = e^{-i\theta}$. The representation results in Theorems 1 and 3 are valid for arbitrary complex numbers provided one replaces $\bar{z}_m$ by z_m^{-1}.

Corresponding to s distinct complex numbers $z_1, \ldots, z_s$ we introduce the polynomials

$$\begin{aligned} p(z) &= \prod_{m=1}^{s}(1 - \bar{z}_m z), \\ p_j(z) &= \prod_{m \neq j}^{s}(1 - \bar{z}_m z) = \tfrac{p(z)}{1-\bar{z}_j z}, \; z \neq z_j, \\ p_{kj}(z) &= \prod_{m \neq k,j}^{s}(1 - \bar{z}_m z) = \tfrac{p(z)}{(1-\bar{z}_k z)(1-\bar{z}_j z)}, \; z \neq z_j, z_k. \end{aligned}$$

2.1 The error correction model

The error correction formulation is a simple consequence of Lagrange's expansion of $A(z)$ around the $s+1$ points $z = 0, z_1, \ldots, z_s$:

Theorem 1 *The polynomial $A(z)$ can be expanded around the points $0, z_1, \ldots, z_s$ as follows*

$$A(z) = p(z)I + \sum_{m=1}^{s} A(z_m)\frac{p_m(z)z}{p_m(z_m)z_m} + p(z)zA_0(z).$$

Proof. The matrix polynomial

$$A(z) - \sum_{m=1}^{s} A(z_m)\frac{p_m(z)z}{p_m(z_m)z_m} - p(z)I$$

is zero for $z = 0, z_1, \ldots, z_s$ and hence each of the entries can be factorized into $p(z)z$ times a polynomial. It follows that the difference can be written as $p(z)zA_0(z)$ for some matrix polynomial $A_0(z)$. ■

An immediate consequence of this is the error correction formulation, see Hylleberg, Engle, Granger and Yoo (1990).

Corollary 2 *If $z_1, \ldots, z_s$ are the roots of $|A(z)| = 0$ then the matrices $A(z_m)$ are of reduced rank such that $A(z_m) = -\alpha_m \beta_m^*$ with α_m and β_m complex matrices of dimension $p \times r_m$ and rank r_m, and X_t satisfies an error correction model*

$$\begin{aligned} p(L)X_t &= \sum_{m=1}^{s} \alpha_m \beta_m^* \tfrac{p_m(L)L}{p_m(z_m)z_m} X_t - p(L)A_0(L)LX_t + \varepsilon_t \\ &= \sum_{m=1}^{s} \alpha_m \beta_m^* X_t^{(m)} - A_0(L)p(L)X_{t-1} + \varepsilon_t, \end{aligned} \tag{2}$$

where we have introduced

$$X_t^{(m)} = \frac{p_m(L)L}{p_m(z_m)z_m} X_t.$$

The idea behind this formulation, see Theorem 4, and condition (4), is that X_t is a non-stationary process and the lag polynomial $p(L)$ makes this process stationary. The processes $X_t^{(m)}, m = 1, \ldots, s$ are non-stationary but, as we shall see below, the components of $X_t^{(m)}$ have the same common non-stationary trend which is removed by the linear combinations β_m^*, and $\beta_m^* X_t^{(m)}$ is stationary, so that the stationary "differences" $p(L)X_t$ react to equilibrium errors given by $\beta_m^* X_t^{(m)}$, through the adjustment coefficients α_m. Note that since the roots may be complex, the coefficients α_m and β_m may be complex, but since the coefficients of $A(z)$ are real, the roots and the coefficients α_m and β_m come in complex conjugate pairs. We use the notation $\beta_m^* = \beta'_{Rm} - i\beta'_{Im}$ for the so-called adjoint matrix. Note also that a different set of roots gives rise to a different error correction formulation.

2.2 Granger's representation theorem

We define the derivative $\dot{A}(z_m)$ of $A(z)$ at $z = z_m$. If the polynomial $|A(z)|$ has a root at $z = z_0$ then $A(z_0)$ is not invertible. We say that $A(z)^{-1}$ has a pole of order k ($k = 0, 1, \ldots$) at $z = z_0$ if

$$\lim_{z \to z_0} (1 - \frac{z}{z_0})^k A^{-1}(z)$$

exists and is non-zero. We next prove a result that gives a necessary and sufficient condition for the inverse function to have a pole of order 1 at a point z_0, say. This condition clearly requires $A(z_0)$ to be singular, but we also need a condition on the derivative of $A(z)$ at z_0, which restricts the behaviour of $A(z)$ in a neighbourhood of z_0. For any (complex) matrix c of dimension $p \times r$ we define $c_\perp$ as a full rank (complex) matrix of dimension $p \times (p - r)$, such that $c' c_\perp = 0$. Note that $(c^*)_\perp = (c_\perp)^*$.

Theorem 3 *If the roots of* $|(A(z)| = 0$ *have the property that* $|z| > 1 + \delta$ *or that* $z \in (z_1, \ldots, z_s)$ *with* $|z_m| = 1$, *then* $A(z_m) = -\alpha_m \beta_m^*$ *and the matrix polynomial* $A(z)$ *is invertible on the disk* $|z| \leq 1 + \delta$ *except at the points* $(z_1, \ldots, z_s)$, *where* $A^{-1}(z)$ *has a pole. A necessary and sufficient condition for the pole to be of order 1 is that*

$$|\alpha_{m\perp}^* \dot{A}(z_m) \beta_{m\perp}| \neq 0, \; m = 1, \ldots, s. \tag{3}$$

In this case we get an expansion of $A^{-1}(z)$ of the form

$$A^{-1}(z) = \sum_{m=1}^{s} C_m \frac{1}{1 - \bar{z}_m z} + C_0(z),\ z \neq (z_1, \ldots, z_s),$$

where

$$\lim_{z \to z_m} (1 - \bar{z}_m z) A(z)^{-1} = C_m = -\bar{z}_m \beta_{m\perp} (\alpha^*_{m\perp} \dot{A}(z_m) \beta_{m\perp})^{-1} \alpha^*_{m\perp},$$

and where $C_0(z)$ has a convergent power series for $|z| \leq 1 + \delta$, that is, $C_0(z)$ has no poles, and $A^{-1}(z)$ has poles of order 1. Moreover it holds that

$$\frac{p_m(z) z}{p_m(z_m) z_m} A^{-1}(z) = C_m \frac{1}{(1 - \bar{z}_m z)} + C_m(z),\ z \neq (z_1, \ldots, z_s)$$

for some power series $C_m(z)$, convergent for $|z| \leq 1 + \delta$.

Proof. The usual expression for the inverse of a matrix

$$A^{-1}(z) = \frac{Adj(A(z))}{|A(z)|},\ z \neq (z_1, \ldots, z_s)$$

shows that $A(z)^{-1}$ has poles at the roots $z = z_1, \ldots, z_s$ since $|A(z_m)| = 0,\ m = 1, \ldots, s$. We want to show that such a pole at $z = z_m$ is of order 1 and has thus the form

$$C_m \frac{1}{(1 - \bar{z}_m z)},$$

such that the continuous function $C_0(z)$ defined by

$$C_0(z) = A(z)^{-1} - \sum_{m=1}^{s} C_m \frac{1}{(1 - \bar{z}_m z)},\ z \neq (z_1, \ldots, z_s),$$

has no poles in the disk $|z| \leq 1 + \delta$, for some $\delta > 0$.

We prove this by investigating the functions in a neighbourhood of each of the poles and show that by subtracting the poles given in the sum we eliminate the poles in $A^{-1}(z)$. Thus we first focus on the root $z = z_1$ where $A(z_1) = -\alpha\beta^*$, and we have left out the subscript to simplify the notation.

Consider therefore a value of z such that $0 < |z - z_1| \leq \varepsilon$. From the expansion

$$A(z) = A(z_1) + (z - z_1)\dot{A}(z_1) + (z - z_1)^2 A_1(z),$$

where $A_1(z)$ is a polynomial, it follows by multiplying by $(\alpha, \alpha_\perp)^*$ and $\left(\beta, \beta_\perp \frac{1}{1 - \bar{z}_1 z}\right)$ that, since $A(z_1) = -\alpha\beta^*$,

$$\begin{aligned}\tilde{A}(z) &= (\alpha, \alpha_\perp)^* A(z)\left(\beta, \beta_\perp \tfrac{1}{1-\bar{z}_1 z}\right) \\ &= \begin{pmatrix} -\alpha^*\alpha\beta^*\beta & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \alpha^* \dot{A}(z_1)\beta(z-z_1) & -z_1\alpha^*\dot{A}(z_1)\beta_\perp \\ \alpha_\perp^*\dot{A}(z_1)\beta(z-z_1) & -z_1\alpha_\perp^*\dot{A}(z_1)\beta_\perp \end{pmatrix} \\ &\quad +(z-z_1)A_2(z),\end{aligned}$$

for some polynomial $A_2(z)$. Here and in the following we often use such a notation for a remainder term, when we expand a polynomial or a power series.

The function $\tilde{A}(z)$ is a matrix polynomial and therefore has no poles. Further

$$\tilde{A}(z_1) = -\begin{pmatrix} \alpha^*\alpha\beta^*\beta & z_1\alpha^*\dot{A}(z_1)\beta_\perp \\ 0 & z_1\alpha_\perp^*\dot{A}(z_1)\beta_\perp \end{pmatrix}$$

has full rank if and only if the assumption (3) holds since

$$|\tilde{A}(z_1)| = (-1)^p|\alpha^*\alpha||\beta^*\beta||z_1\alpha_\perp^*\dot{A}(z_1)\beta_\perp| \neq 0.$$

In this case $\tilde{A}(z)$ is invertible for $|z - z_1| \leq \varepsilon$ for some ε, such that $0 < \varepsilon < \min_{k\neq l} |z_k - z_l|$, and we find by the expansion

$$\tilde{A}^{-1}(z) = \tilde{A}^{-1}(z_1) + (z-z_1)M_2(z),$$

that for $z \neq z_1$

$$\begin{aligned}A^{-1}(z) &= \left(\beta, \beta_\perp \tfrac{1}{1-\bar{z}_1 z}\right)\tilde{A}^{-1}(z)(\alpha,\alpha_\perp)^* \\ &= -\tfrac{\bar{z}_1}{1-\bar{z}_1 z}\beta_\perp(\alpha_\perp^*\dot{A}(z_1)\beta_\perp)^{-1}\alpha_\perp^* + M_3(z) \\ &= C_1\tfrac{1}{1-\bar{z}_1 z} + M_3(z)\end{aligned}$$

where C_1 is given in Theorem 3 and

$$M_3(z) = ((z-z_1)\beta, -\beta_\perp z_1)M_2(z)(\alpha,\alpha_\perp)^*.$$

Here $M_2(z)$ and $M_3(z)$ are convergent power series which are notations for the remainder terms in the expansions. Hence $A^{-1}(z) - C_1\frac{1}{1-\bar{z}_1 z}$ has no pole at $z = z_1$ and extends by continuity to the point $z = z_1$.

The same argument can be used to remove the other poles from $A^{-1}(z)$ and the theorem has been proved. ■

The next result is a representation of the solution of the error correction model (2).

Theorem 4 *Let the equation $|A(z)| = 0$ have roots outside the unit disk and at $z_1, \ldots, z_s$ with absolute value 1. Let for $m = 1, \ldots, s$*

$$A(z_m) = -\alpha_m \beta_m^*,$$

where α_m and β_m are $p \times r_m$ of rank $r_m < p$, and assume that

$$|\alpha_{m\perp}^* \dot{A}(z_m) \beta_{m\perp}| \neq 0. \tag{4}$$

Let X_t be the solution of the error correction model (2), then X_t is non-stationary and $p(L)X_t$ and $p_m(L)\beta_m^* X_t$ can be made stationary by a suitable choice of initial distribution. In this case the processes X_t and $X_t^{(m)}$ can be given the representation

$$X_t = \sum_{m=1}^{s} C_m \bar{z}_m^t S_t^{(m)} + \sum_{m=1}^{s} \bar{z}_m^t A_m + Y_t, \tag{5}$$

and

$$X_t^{(m)} = \frac{p_m(L)L}{p_m(z_m)z_m} X_t = C_m \bar{z}_m^t S_{t-1}^{(m)} + \bar{z}_m^t A_m + Y_t^{(m)}, \tag{6}$$

where $S_t^{(m)}$ is given by

$$S_t^{(m)} = \sum_{j=0}^{t} z_m^j \varepsilon_j,$$

and the random variables A_m depends on initial conditions such that $\beta_m^* A_m = 0$, and finally Y_t and $Y_t^{(m)}$ are stationary processes.

Thus the non-stationary process X_t can be made stationary by the difference operator $p(L)$ and, since $\beta_m^* X_t^{(m)}$ is stationary, we call X_t *seasonally cointegrated at $z_m = e^{i\theta_m}$, or at frequency θ_m, with cointegrating vectors β_m*. Note that $S_t^{(m)}$ is not a random walk since $\Delta S_t^{(m)} = z_m^t \varepsilon_t$ are independent but not in general identically distributed. We call such a process a seasonal random walk. Note also the factor $\bar{z}_m^t$ in front of $S_t^{(m)}$ gives a type of non-stationarity that is different from the usual unit root non-stationarity. Finally note that since we allow for complex roots, the process $X_t^{(m)}$ and the coefficients α_m and β_m are in general complex. Since, however, the data and the coefficients in $A(z)$ are real, the roots come in complex conjugate pairs and hence a reduced rank condition at a complex root automatically implies a reduced rank condition at the complex conjugate root. This will complicate the statistical analysis below.

The difference between the results in Theorem 3 and Theorem 4 is that in order to interpret Theorem 3 for stochastic processes, care has to be taken of initial values in the representation (5) and (6) in order to translate the results about the power series into results about the lag operator.

Proof. From Theorem 3 we find

$$p(z)A(z)^{-1} = \sum_{m=1}^{s} C_m p_m(z) + p(z)C_0(z).$$

Expressed in terms of the lag operator L, defined by $LX_t = X_{t-1}$, we get

$$p(L)X_t = \sum_{m=1}^{s} C_m p_m(L)\varepsilon_t + p(L)Y_t, \tag{7}$$

where $Y_t = C_0(L)\varepsilon_t$. Since the right hand side is stationary this representation shows that we can choose the initial values of the process X_t such that the process $p(L)X_t$ becomes stationary.

We want to solve this equation for X_t by removing the polynomial $p(L)$ one factor at a time by summation.

Consider first the root $z = z_1$. The definition of $p(z)$ implies that

$$p(z) = (1 - \bar{z}_1 z)p_1(z),\ p_m(z) = (1 - \bar{z}_1 z)p_{m1}(z),\ m \neq 1,$$

and we can write equation (7) as

$$(1 - \bar{z}_1 L)\left[p_1(L)(X_t - Y_t) - \sum_{m=2}^{s} C_m p_{m1}(L)\varepsilon_t\right] = C_1 p_1(L)\varepsilon_t.$$

Solving these equations we find

$$p_1(L)(X_t - Y_t) - \sum_{m=2}^{s} C_m p_{m1}(L)\varepsilon_t = \bar{z}_1^{t+1}A + C_1 p_1(L)\bar{z}_1^t \sum_{m=0}^{t} z_1^m \varepsilon_m. \tag{8}$$

Here A is the initial value of the left hand side. Next notice that

$$p_1(L)\bar{z}_1^t = p_1(z_1)\bar{z}_1^t,$$

such that result of the above calculations can be expressed as

$$p_1(L)(X_t - Y_t - C_1\bar{z}_1^t S_t^{(1)} - \bar{z}_1^t A_1) = \sum_{m=2}^{s} C_m p_{m1}(L)\varepsilon_t, \tag{9}$$

with $A_1 = \frac{\bar{z}_1 A}{p_1(z_1)}$. If we choose the initial values such that $\beta_1^* A_1 = 0$, then we get the equation

$$p_1(L)\beta_1^* X_t = p_1(L)\beta_1^* Y_t + \sum_{m=2}^{s} \beta_1^* C_m p_{m1}(L)\varepsilon_t,\ t = 1, \ldots, T,$$

since $\beta_1^* C_1 = 0$. This shows that, with this choice of initial values, we can represent $p_1(L)\beta_1^* X_t$ as a stationary process.

Equation (9) has exactly the same form as the one we started with in (7) except that the root $z = z_1$ has been removed. In the same way we can successively eliminate the roots by summation of the equation each time subtracting a term of the form $C_m \bar{z}_m^t S_t^{(m)} + \bar{z}_m^t A_m$ from the left hand side and thereby prove the representation. ■

We next apply Theorem 3 to solve the autoregressive equations also in the case where they contain deterministic terms.

Thus assume that X_t is the solution to the equations

$$A(L)X_t + \Phi D_t = \epsilon_t,$$

where $A(z)$ satisfies the conditions set out in Theorem 3. In this case one gets

$$p(L)X_t = \sum_{m=1}^{s} C_m p_m(L)(\epsilon_t + \Phi D_t) + p(L)C_0(L)(\epsilon_t + \Phi D_t),$$

which can be solved for X_t to give

$$X_t = \sum_{m=1}^{s} C_m \bar{z}_m^t S_t^{(m)} + \sum_{m=1}^{s} C_m \bar{z}_m^t \Phi \sum_{j=0}^{t} z_m^j D_j + \sum_{m=1}^{s} \bar{z}_m^t \tilde{A}_m + \tilde{Y}_t,$$

where $\tilde{Y}_t - E(\tilde{Y}_t)$ is stationary, and $\tilde{A}_m$ depends on initial values such that $\beta_m^* \tilde{A}_m = 0$. It is seen that if $D_j = 1$, say, then the deterministic term gives rise to the term $\bar{z}_m^t \sum_{j=0}^t z_m^j = \bar{z}_m^t \frac{1 - z_m^{t+1}}{1 - z_m}$, which remains bounded unless $z_m = 1$ in which case we get a linear trend in the process. Similarly if $D_t = t$ we find

$$\bar{z}_m^t \sum_{j=0}^{t} j z_m^j = \frac{t+1}{1 - \bar{z}_m} - \frac{(1 - \bar{z}_m^{t+1})}{(1 - \bar{z}_m)^2},$$

which is a linear trend if $z_m \neq 1$ and a quadratic trend if $z_m = 1$. Thus if $\Phi D_t = \Phi_0 + \Phi_1 t$ we get a quadratic trend which vanishes, if we choose $\alpha_{1\perp}' \Phi_1 = 0$. If $\Phi_1 = 0$, we get a linear trend which vanishes when $\alpha_{1\perp}' \Phi_0 = 0$. The next subsection discusses the same result for seasonal dummies.

2.3 The role of seasonal dummies

A special case of a deterministic term is when the data is measured with a given frequency $\tilde{s}$, say, per year, and we include seasonal dummies to pick up the changing mean in the process. We want to consider the case where the unit roots of the process are among the roots of unity corresponding to $\tilde{s}$, that is, of the form $e^{\frac{2\pi i m}{\tilde{s}}}$, $m = 1, \ldots, \tilde{s}$. This is the situation if for instance we have quarterly data and a root in the process at $z = 1$. We denote the roots of unity z_m, $m = 1, \ldots, \tilde{s}$ and assume for simplicity here that $z_1 = 1$, $z_2, \ldots, z_s$ are roots of the process, $s \leq \tilde{s}$. The results are easily modified if this is not the case.

The seasonal dummies are defined by the p vector D_t, with the property that $D_t = D_{t+\tilde{s}}$. If we consider this as a difference equation the characteristic roots are exactly the roots of unity, and we can express the solution as

$$D_t = \sum_{m=1}^{\tilde{s}} \bar{z}_m^t d_m,$$

for some (real) linearly independent vectors d_m which can be determined by the initial $\tilde{s}$ values of D_t

$$d_m = \frac{1}{\tilde{s}} \sum_{j=1}^{\tilde{s}} \bar{z}_m^j D_j.$$

With this notation we find that

$$\sum_{j=0}^{t} z_m^j D_j = \sum_{j=0}^{t} z_m^j \sum_{n=1}^{\tilde{s}} \bar{z}_n^j d_n = t d_m + \sum_{n \neq m} \frac{1 - z_m^{t+1} \bar{z}_n^{t+1}}{1 - z_m \bar{z}_n} d_n, \; m = 1, \ldots, s,$$

which shows that the seasonal dummy generates a trend in the process

$$\bar{z}_m^t t$$

with coefficient $C_m \Phi d_m$ at the unit root z_m, $m = 1, \ldots, s$. For $z_m \neq 1$ this trend has an oscillating behaviour due to the factor $\bar{z}_m^t$, which is unwanted in the description of data. We can remove the trend by assuming the restriction on the parameters

$$C_m \Phi d_m = 0, \text{ or } \alpha_{m\perp}^* \Phi d_m = 0.$$

This result was first proved by Kunst and Franses (1995).

We reparametrize the model by introducing the parameters $\Phi_m = \Phi d_m$. The vectors d_m come in complex conjugate pairs, which also holds for the new parameters. The deterministic terms in the equation becomes

$$\Phi D_t = \sum_{m=1}^{\tilde{s}} \Phi_m z_m^t.$$

We now restrict $\Phi_m, m = 2, \ldots, s$ by $\alpha^{*\prime}_{m\perp}\Phi_m = 0$, or $\Phi_m = \alpha_m \rho^{*}_m$. In this way we avoid the oscillating trends but leave the possibility open of a linear trend generated by the unit root $z = 1$. If we also want to restrict this we further assume that $\alpha'_{1\perp}\Phi_1 = 0$.

We conclude this section by a few illustrative examples.

2.4 Examples

We give here various simple examples for models for annual, semi-annual and quarterly data.

2.4.1 Annual data

If $z = 1$ is the only unit root in the process, then $p(z) = 1 - z$ and $X_t^{(1)} = X_{t-1}$, and we get that (2) reduces to the usual error correction model for $I(1)$ variables

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t,$$

where we have left out further dynamics and deterministic terms.

2.4.2 Semi-annual data

If the unit roots are $z = \pm 1$ in the process, then $p(z) = (1 - z)(1 + z) = 1 - z^2$ and we find

$$\begin{aligned} X_t^{(1)} &= \frac{(1+L)L}{2} X_t = \frac{1}{2}(X_{t-1} + X_{t-2}), \\ X_t^{(2)} &= \frac{(1-L)L}{2} X_t = \frac{1}{2}(X_{t-1} - X_{t-2}), \end{aligned}$$

such that (2) becomes

$$X_t - X_{t-2} = \frac{1}{2}\alpha_1\beta_1'(X_{t-1} + X_{t-2}) + \frac{1}{2}\alpha_2\beta_2'(X_{t-1} - X_{t-2}) + \varepsilon_t.$$

In this simple case we can interpret the results. Consider for instance a process consisting of semi-annual income and consumption. In this case $X_t^{(1)}$ is just the annual average, and the model specifies that this process is a non-stationary $I(1)$ process, which cointegrates, such that annual consumption follows annual income in a stationary way through the cointegrating coefficients β_1. The process $X_t^{(2)}$, however, measures the variation within a year and the models specifies

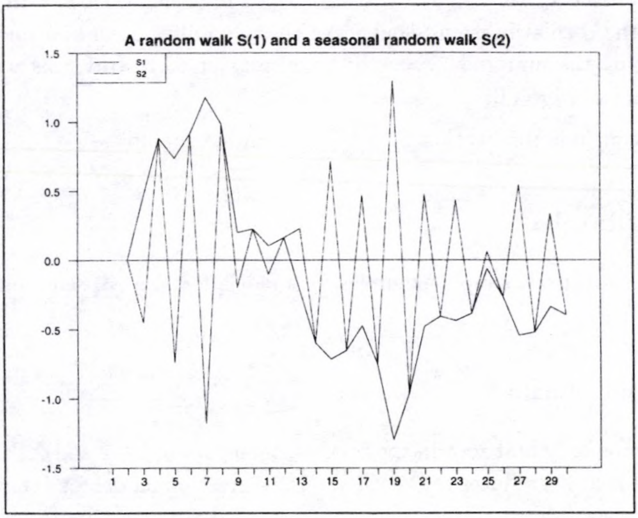

Figure 1:

that such a process has a seasonal non-stationarity, which means that when averaged within a year it becomes stationary. The cointegration vector β_2 specifies that linear combination of the annual variation of consumption and income that cointegrates.

Thus not only the non-stationary yearly averages but also the non-stationary seasonal variation within a year has to move together according to the model.

In order to understand the type of non-stationarity induced by a unit root at $z = -1$, consider the process

$$X_t = (-1)^t S_t^{(2)} = (-1)^t \sum_{j=0}^{t} (-1)^j \varepsilon_j = \sum_{j=0}^{t} (-1)^j \varepsilon_{t-j},$$

which enters the representation theorem. In Figure 1 we have generated $\varepsilon_1, \ldots, \varepsilon_T$ i.i.d. $N(0,1), T = 30$.

Since the normal distribution is symmetric, the process $S_t^{(2)}$ is a random walk, and the factor $(-1)^t$ changes the sign of every second term, which give rise to the oscillating behaviour that we see in seasonally varying processes. It is obvious from the picture that differencing such a process will not give stationarity, whereas one can obtain a stationary process by smoothing using a moving average.

Note that when the random walk is positive for an interval then X_t oscillates systematically between positive and negative values, but when $S_t^{(2)}$ gets too close to zero, or a large draw of ε_t occurs, then it can change sign with the result that the peaks of X_t are shifted one period, such that "summer becomes winter". This property is an intrinsic property of processes generated by the error correction model allowing for seasonal cointegration. It is easy in the example to check the role of constant, linear term and seasonal dummies on the behaviour of the process.

2.4.3 Quarterly data

Next we consider the situation where we have quarterly data and unit roots in the process at $z = \pm 1, \pm i$. In this case have

$$p(z) = (1-z)(1+z)(1+iz)(1-iz) = (1-z)^4.$$

The processes that are needed in the error correction model are

$$\begin{aligned}
X_t^{(1)} &= \tfrac{1}{4}(X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4}),\\
X_t^{(2)} &= \tfrac{1}{4}(X_{t-1} - X_{t-2} + X_{t-3} - X_{t-4}),\\
X_t^{(3)} &= \tfrac{1}{4i}(X_{t-1} + iX_{t-2} - X_{t-3} - iX_{t-4}),\\
X_t^{(4)} &= -\tfrac{1}{4i}(X_{t-1} - iX_{t-2} - X_{t-3} + iX_{t-4}).
\end{aligned}$$

The error correction model contains 4 terms, and we would like to express them using real variables, see also (15). If we let $X_t^{(3)} = X_{Rt}^{(3)} + iX_{It}^{(3)}$ and $X_t^{(4)} = \bar{X}_t^{(3)} = X_{Rt}^{(3)} - iX_{It}^{(3)}$ we find

$$\begin{aligned}
X_{Rt}^{(3)} &= \frac{1}{4}(X_{t-2} - X_{t-4}),\\
X_{It}^{(3)} &= -\frac{1}{4}(X_{t-1} - X_{t-3}).
\end{aligned}$$

The error correction model then becomes

$$\begin{aligned}
X_t - X_{t-4} &= \alpha_1\beta_1' X_t^{(1)} + \alpha_2\beta_2' X_t^{(2)}\\
&\quad +(\alpha_R\beta_R' + \alpha_I\beta_I')(X_{t-2} - X_{t-4})\\
&\quad +(\alpha_R\beta_I' - \alpha_I\beta_R')(X_{t-1} - X_{t-3}) + \varepsilon_t,
\end{aligned}$$

where we have absorbed the factor $\frac{1}{4}$ into the coefficients, and for ease of notation we let $\alpha_3 = \alpha_R + i\alpha_I$, $\alpha_4 = \bar{\alpha}_3$, $\beta_3 = \beta_R + i\beta_I$, $\beta_4 = \bar{\beta}_3$.

Note that the coefficient matrix to $X_{t-2} - X_{t-4}$ is rather complicated. It need not even have reduced rank. The same parameters appear in the coefficient to $X_{t-1} - X_{t-3}$. It has been suggested to assume that $\alpha_R\beta_I' - \alpha_I\beta_R' = 0$, in order simplify the equations, but it is seen that this is a peculiar restriction on all the coefficients, which is hard to interpret. If we assume instead that $\beta_I = 0$ we get some simplification and the equations contain the term

$$\alpha_R\beta_R'(X_{t-2} - X_{t-4}) - \alpha_I\beta_R'(X_{t-1} - X_{t-3}).$$

This has the advantage that only one linear combination of $(1 - L^2)X_t$ appears, and the interpretation is that $\beta_{3R}'(X_{t-2} - X_{t-4})$ is either stationary or cointegrates with its own lag. Thus we have a type of multicointegration. If we also assume that $\alpha_I = 0$ (or $\alpha_R = 0$) we get the simple result that the equation contains a term of the form $\alpha_R\beta_R'(X_{t-2} - X_{t-4})$ and the interpretation that the process $X_{t-2} - X_{t-4}$ cointegrates.

We get a different error correction model if the process only has roots at $z = 1, \pm i$, since then $p(z) = (1 - z)(1 + iz)(1 - iz) = (1 - z)(1 + z^2)$. In this case

$$\begin{aligned}
X_t^{(1)} &= \frac{(1 + L^2)}{2} L X_t = \frac{1}{2}(X_{t-1} + X_{t-3}),\\
X_t^{(2)} &= \frac{(1 - L)(1 - iL)}{(1 - i)2} L X_t = \frac{1}{2(1 - i)}(\Delta X_{t-1} - i\Delta X_{t-2}),\\
X_t^{(3)} &= \frac{(1 - L)(1 + iL)}{(1 + i)2} L X_t = \frac{1}{2(1 + i)}(\Delta X_{t-1} + i\Delta X_{t-2}).
\end{aligned}$$

We find the real and imaginary part as follows:

$$X_{Rt}^{(2)} = \frac{1}{4}\Delta_2 X_{t-1},\ X_{It}^{(2)} = \frac{1}{4}\Delta^2 X_{t-1}.$$

The error correction model becomes

$$\begin{aligned}
& X_t - X_{t-1} + X_{t-2} - X_{t-3}\\
= \ & \alpha_1\beta_1'(X_{t-1} + X_{t-3}) + (\alpha_R\beta_R' + \alpha_I\beta_I')\Delta_2 X_{t-1} + (\alpha_R\beta_I' - \alpha_I\beta_R')\Delta^2 X_{t-1} + \varepsilon_t.
\end{aligned}$$

In this section we have given a general version of the Granger Representation Theorem which clarifies when we get a seasonally cointegrated solution to the autoregressive equations and when the solution is integrated of order 1 at the seasonal frequency. We also gave a discussion of the trends generated by constant, linear term and seasonal dummies, and what restrictions are necessary in order to avoid them, if the unit roots of the process are among the roots of unity corresponding to the frequency of the data. In the next section we use these results to define the statistical model we want to analyze.

3 The model for seasonal cointegration and its statistical analysis

In this section we define the statistical model for autoregressive processes of order 1 at seasonal frequency which allows for seasonal cointegration. We give various models defined by restrictions on the deterministic terms. We discuss the Gaussian maximum likelihood estimation and the formulation of some hypotheses on the cointegrating ranks and some interesting hypotheses on the cointegrating vectors and the adjustment coefficients.

3.1 The statistical models defined by restrictions on the deterministic terms

The $p-$dimensional vector autoregressive model for seasonal cointegration is defined by the equations

$$p(L)X_t = \sum_{m=1}^{s} \alpha_m \beta_m^* X_t^{(m)} + \sum_{j=1}^{k} \Gamma_j p(L) X_{t-j} + \Phi D_t + \varepsilon_t, \; t = 1, \ldots, T. \quad (10)$$

Here ε_t are i.i.d. $N_p(0, \Omega)$, and the parameters α_m, β_m, $m = 1, \ldots, s$, Γ_j, $j = 1, \ldots, k$, Φ and Ω are freely varying, except that the α_m and β_m come in complex conjugate pairs. We assume that D_t consists of deterministic terms. Note that the lag length is $l = k + s$. The dimension of α_m and β_m is $p \times r_m$, and the initial values are fixed.

If in particular the roots of the process are also roots of unity, corresponding to a given frequency $\tilde{s}$ of the data, we can introduce seasonal dummies D_t in the model. As seen in Section 2 they give rise to trends in the process and it was shown how these can be avoided by restriction of the parameters. We decompose the parameter Φ as $\Phi D_t = \sum_{m=1}^{\tilde{s}} \Phi_m z_m^t$.

We give here the model where we assume that there are no trends in the process so that $\Phi_m = \alpha_m \rho_m^*$, for some $\rho_m \times 1$ matrix ρ_m, $m = 1, \ldots, s$, which imply that

$$p(L)X_t = \sum_{m=1}^{s} \alpha_m \begin{pmatrix} \beta_m \\ \rho_m \end{pmatrix}^* \begin{pmatrix} X_t^{(m)} \\ z_m^t \end{pmatrix} + \sum_{j=1}^{k} \Gamma_j p(L) X_{t-j} + \sum_{m=s+1}^{\bar{s}} \Phi_m z_m^t + \varepsilon_t. \tag{11}$$

If we want to allow for a linear trend we do not restrict at zero frequency but use the model

$$p(L)X_t = \alpha_1 \beta_1' X_t^{(1)} + \sum_{m=2}^{s} \alpha_m \begin{pmatrix} \beta_m \\ \rho_m \end{pmatrix}^* \begin{pmatrix} X_t^{(m)} \\ z_m^t \end{pmatrix} + \sum_{j=1}^{k} \Gamma_j p(L) X_{t-j} + \Phi_1 + \sum_{m=s+1}^{\bar{s}} \Phi_m z_m^t + \varepsilon_t. \tag{12}$$

In these two last models the parameters specified are varying freely, with the only restriction that the parameters α_m, β_m and Φ_m come in complex conjugate pairs. Note how the roots of unity that do not correspond to unit roots in the process enter with pairwise unrestricted coefficients and do not give rise to trends in the process.

3.2 Some algorithms for the estimation

The statistical analysis of (10) leads to a non-linear regression problem since the coefficients α_m and β_m enter through their product. We here discuss estimation of the model without restrictions on the deterministic term and mention in the end of this subsection how to modify the algorithm if the deterministic terms are restricted, as in models (11) and (12).

Since the cointegration model (10) does not restrict the matrices Γ_j and Φ we can concentrate the likelihood function with respect to these and define the residuals R_{0t}, $R_{1t}^{(m)}$ and $R_{\varepsilon t}$ by regression of $p(L)X_t$, $X_t^{(m)}$ and ε_t on D_t and lagged values of $p(L)X_t$. Thus we get the equation

$$R_{0t} = \sum_{m=1}^{s} \alpha_m \beta_m^* R_{1t}^{(m)} + R_{\varepsilon t}. \tag{13}$$

An algorithm for estimating this model, see Boswijk (1995), can be found by noting that for fixed β coefficients the model is a linear regression model that determines the $\alpha's$ and Ω by simple regression of p variables R_{0t} on $\sum_{m=1}^{s} r_m$ variables $\beta_1^* R_{1t}^{(1)}, \ldots, \beta_m^* R_{1t}^{(m)}$. For fixed values of α and Ω, however, we have a linear regression model in the β coefficients which can be estimated by generalized least squares. This determines a switching algorithm, which in each step increases the likelihood function. but the second step involves vectorizing β_m so we need a total of $p \sum_{m=1}^{s} r_m$ regressors.

Another algorithm can be based on the first and second derivatives of the likelihood function and an application of the Gauss Newton algorithm. This algorithm also involves a large number of variables in general.

Finally we describe an algorithm which is slightly simpler, and which can be proved to give estimators which are asymptotically equivalent to the maximum likelihood estimators, since the regressors $X_t^{(m)}$ are asymptotically uncorrelated, in the sense that

$$T^{-1} \sum_{t=1}^{T} X_t^{(m)} X_t^{(n)*} \xrightarrow{P} 0, \; z_n \neq z_m,$$

see Corollary 7.

The idea of the algorithm is that when focussing on one frequency we can concentrate out the other regressors by ignoring the constraint of reduced rank at these other frequencies, see Lee (1992). We illustrate the situation of a complex root, since the real roots 1 and -1 are easily handled the same way.

Consider therefore the situation where, say, $z_1 = e^{i\theta}$ and $z_2 = e^{-i\theta}$ are two complex roots with $0 < \theta < \pi$. Note that $A(e^{i\theta}) = \bar{A}(e^{-i\theta})$ and $X_t^{(2)} = \bar{X}_1^{(1)}$. For notational reasons we use α and β without subscripts now and let $\alpha_1 = \alpha$, $\alpha_2 = \bar{\alpha}$, $\beta_1 = \beta$, and $\beta_2 = \bar{\beta}$. Thus we write the model equation (13) as

$$R_{0t} = \alpha\beta^* R_{1t}^{(1)} + \bar{\alpha}\bar{\beta}^* \bar{R}_{1t}^{(1)} + \sum_{m=3}^{s} \alpha_m \beta_m^* R_{1t}^{(m)} + R_{\varepsilon t}. \tag{14}$$

We concentrate with respect to $R_{1t}^{(m)}$ where $m \neq (1,2)$, that is, we remove the restriction of reduced rank at $z_3, \ldots, z_s$. This gives residuals U_{0t}, U_{1t}, and

$U_{\varepsilon t}$ and we find the equations

$$\begin{aligned}
U_{0t} &= \alpha\beta^* U_{1t} + \bar{\alpha}\bar{\beta}^* \bar{U}_{1t} + U_{\varepsilon t} \\
&= 2Real[(\alpha_R + i\alpha_I)(\beta_R - i\beta_I)'(U_{Rt} + iU_{It})] + U_{\varepsilon t} \\
&= 2[(\alpha_R\beta_R' + \alpha_I\beta_I')U_{Rt} + (\alpha_R\beta_I' - \alpha_I\beta_R')U_{It}] + U_{\varepsilon t} \\
&= 2(\alpha_R, -\alpha_I)\begin{pmatrix} \beta_R & -\beta_I \\ \beta_I & \beta_R \end{pmatrix}'\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix} + U_{\varepsilon t} \\
&= \tilde{\alpha}\boldsymbol{\beta}'\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix} + U_{\varepsilon t},
\end{aligned} \tag{15}$$

where we use the notation

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_R & -\beta_I \\ \beta_I & \beta_R \end{pmatrix}, \tilde{\alpha} = 2(\alpha_R, -\alpha_I), \boldsymbol{\alpha} = \begin{pmatrix} \alpha_R & -\alpha_I \\ \alpha_I & \alpha_R \end{pmatrix}.$$

In Appendix A the matrix representation of complex numbers and matrices is explained. This representation is noted throughout by boldface. Since the roots come in complex pairs the sum $\sum_{m=3}^{s} \alpha_m \beta_m^* X_t^{(m)}$ is real, such that both U_{0t} and $U_{\varepsilon t}$ are real. The statistical problem appears to be a reduced rank regression problem, at least if $p > 2r$, but the matrix $\boldsymbol{\beta}$ is not unrestricted but has complex structure.

In order to express the partially maximized likelihood function we introduce the product moments

$$\begin{aligned}
S_{11} &= T^{-1}\sum_{t=1}^{T}\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix}\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix}', \ (2p \times 2p), \\
S_{10} &= T^{-1}\sum_{t=1}^{T}\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix} U_{0t}', \ (2p \times p), \\
S_{00} &= T^{-1}\sum_{t=1}^{T} U_{0t}U_{0t}', \ (p \times p), \\
S_{\varepsilon 1} &= T^{-1}\sum_{t=1}^{T} U_{\varepsilon t}\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix}', \ (p \times 2p).
\end{aligned}$$

Finally we define $S_{11.0} = S_{11} - S_{10}S_{00}^{-1}S_{01}$.

For fixed value of $\boldsymbol{\beta}$ we can concentrate the likelihood function with respect to the parameters $\tilde{\alpha} = 2(\alpha_R, -\alpha_I)$ and Ω and find, apart from a constant factor,

$$L_{\max}^{-\frac{2}{T}}(\boldsymbol{\beta}) = |\hat{\Omega}| = |S_{00} - S_{01}\boldsymbol{\beta}(\boldsymbol{\beta}'S_{11}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'S_{10}| = |S_{00}|\frac{|\boldsymbol{\beta}'S_{11.0}\boldsymbol{\beta}|}{|\boldsymbol{\beta}'S_{11}\boldsymbol{\beta}|}. \tag{16}$$

This minimization cannot be solved as an eigenvalue problem since the $2p \times 2r$ matrix $\boldsymbol{\beta}$ has complex structure while S_{11} and $S_{11.0}$ do not have complex structure.

We can minimize (16) by an iterative procedure using the Gauss Newton algorithm or we can use the idea of switching between $(\tilde{\alpha}, \Omega)$ and $\boldsymbol{\beta}$ in (15). Applying the switching algorithm here only involves $2pr_m$ parameters from the cointegrating relations and a similar number from the adjustment coefficients.

Finally the maximum likelihood estimator can be calculated iteratively as follows. For fixed values of $\beta_2, \ldots, \beta_s$ we can concentrate the likelihood function with respect to $\alpha_2, \ldots, \alpha_s$. Then the equations have the form (15) and we can apply the switching algorithm to determine α_1 and β_1. One can save time by not switching to convergence. Next fix $\beta_1, \beta_3, \ldots, \beta_s$ and repeat the procedure as above. In this way one can by focussing on one frequency at a time reduce the dimension of the matrices involved in the regressions.

If instead we consider the problem of reduced rank at $\theta = 0$ or π then we get the product moments as before but now with $X_t^{(1)}$, say, corrected for all the other components. In this case all residuals are real and the matrices S_{11}, S_{01}, and S_{00} are all of dimension $p \times p$, and the problem can then be solved by reduced rank regression, see Lee (1992).

Finally we can use the same ideas to estimate the models (11) and (12) with the various restrictions on the deterministic terms. The coefficients Φ_j with $j \geq s$, and possibly $j = 1$, can be concentrated out in the preliminary regression, and in the reduced rank regressions we just replace $X_t^{(m)}$ by the extended variables $(X_t^{(m)\prime}, z_m^t)'$. Corresponding to the equation (15) we get

$$\begin{aligned}
\Phi_m z_m^t + \bar{\Phi}_m \bar{z}_m^t &= 2\,\mathrm{Re}(\Phi_m z_m^t) = 2\,\mathrm{Re}(\Phi_{Rm} + i\Phi_{mI})(\cos(\theta_m t) + i\sin(\theta_m t)) \\
&= 2(\Phi_{Rm}\cos(\theta_m t) + \Phi_{mI}\sin(\theta_m t)),
\end{aligned}$$

and

$$\rho_m z_m^t + \bar{\rho}_m \bar{z}_m^t = 2(\rho_{Rm}\cos(\theta_m t) + \rho_{mI}\sin(\theta_m t)).$$

In this case we define the regressors derived from $U_{1t}^{(m)}, \theta_m$ as

$$(U_{Rt}^{(m)\prime}, \cos(\theta_m t), U_{It}^{(m)\prime}, \sin(\theta_m t))'$$

and the cointegrating coefficient is

$$\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\rho} \end{pmatrix} = \begin{pmatrix} \beta_{mR} & -\beta_{mI} \\ \rho_{mR} & -\rho_{mI} \\ \beta_{mI} & \beta_{mR} \\ \rho_{mI} & \rho_{mR} \end{pmatrix}.$$

This shows how, in model (11) we can concentrate out the coefficients $\Phi_j,\ j > s$, using simple regression. The switching algorithm is then applied to the extended variables, where the residuals are extended by a cos or a sin. In model (12) we eliminate Φ_1 by regression and do not extend the variable $X_t^{(1)}$.

The algorithm has been programmed in Gauss, see Schaumburg (1996) and RATS, see Dahl Pedersen (1996).

3.3 Hypotheses of interest

The main hypothesis of interest is of course the test for reduced rank at the various complex frequencies. This requires maximization of the likelihood function under model $H(r)$, that is, the assumption of reduced rank r at the complex frequency θ as discussed in the previous subsection. We then compare the obtained maximum with the maximum obtained from the unrestricted VAR, which corresponds to $r = p$. Thus the test statistic is

$$-2\log Q(H(r)|H(p)) = 2\log\left(\frac{|\boldsymbol{\beta}' S_{11.0}\boldsymbol{\beta}||S_{11}|}{|\boldsymbol{\beta}' S_{11}\boldsymbol{\beta}||S_{11.0}|}\right).$$

Other hypotheses of interest are hypotheses on the cointegrating coefficients $\boldsymbol{\beta}$. The most interesting perhaps is the hypothesis that $\boldsymbol{\beta}$ is real, since without this simple structure the interpretation becomes rather tedious. This hypothesis is formulated by Lee (1992), and in the present notation becomes the restriction

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_R & 0 \\ 0 & \beta_R \end{pmatrix}. \tag{17}$$

Due to the non-identification of $\boldsymbol{\beta}$ we can give an equivalent formulation of the hypothesis as $\beta_R = \beta_I$. Finally we can consider the assumption that

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_R & 0 \\ 0 & \alpha_R \end{pmatrix}, \tag{18}$$

which allows for a simple interpretation, see the examples in sub-section 2.4. The maximization of the concentrated likelihood function (16) under any of these restrictions again requires an iterative algorithm for finding the maximum. By comparing the obtained maxima with and without the restrictions (17) and (18) we obtain the likelihood ratio test statistic.

Clearly if there are prior hypotheses about the structure of the cointegrating relations we can test those by the likelihood ratio test, by suitably modifying the algorithm for finding the maximum.

The asymptotic results that allow these procedures to be used are given in the next sections.

4 Asymptotic results

This section deals with some technical results on asymptotic behaviour of various processes and product moments. The proofs are given in the Appendix A. We assume throughout that the processes are generated by the autoregressive equations without deterministic terms and that the ε are i.i.d. with mean zero and variance Ω. We start with the sums $S_t^{(m)}$ and then find the limiting behaviour of $X_t^{(m)}$ and finally investigate S_{11}, S_{10}, S_{00} and $S_{\varepsilon 1}$ which are based on the residuals from the regression (15).

The limit distribution of the random walk $S_t^{(m)}$ is found in Chan and Wei (1988) who show the following result:

Lemma 5 *If* $S_t^{(m)} = \sum_{j=0}^t z_m^j \varepsilon_j$ *and* $z_m = \exp(i\theta_m)$, *then*

$$T^{-\frac{1}{2}} S_{[Tu]}^{(m)} \xrightarrow{w} \delta_m \left(W_R^{(m)}(u) + iW_I^{(m)}(u) \right) = \delta_m W_m, \quad \delta_m = \begin{cases} \frac{1}{\sqrt{2}} & 0 < \theta_m < \pi \\ 1 & \theta_m = 0,\, \pi \end{cases} \tag{19}$$

where $W_R^{(m)}$ *and* $W_I^{(m)}$ *are independent Brownian motions with variance matrix* Ω. *Moreover these Brownian motions are independent for different values of* θ_m. ∎

Another result that follows from their calculations is the following.

Theorem 6 *For* $T \to \infty$

$$T^{-2} \sum_{i=1}^T S_t^{(m)} S_t^{(n)*} \xrightarrow{w} \delta_m \delta_n \int_0^1 W_m W_n^* du, \tag{20}$$

$$T^{-1} \sum_{i=1}^T S_{t-1}^{(m)} \bar{z}_n^t \varepsilon_t' \xrightarrow{w} \delta_m \delta_n \int_0^1 W_m (dW_n^*). \tag{21}$$

If further $f(t)$ *is a matrix valued function such that* $F(t) = \sum_{i=1}^t f(i)$ *is bounded, then*

$$T^{-2} \sum_{i=1}^T S_t^{(m)} f(t) S_t^{(n)*} \xrightarrow{P} 0. \tag{22}$$

■

In the study of the asymptotic properties of the product moments that appear in the statistical analysis we need to know the joint behaviour of all the random walks that come from the various components $X_t^{(m)}$. From Theorem 6 one obtains

Corollary 7 *The asymptotic properties of the product moment matrices are given by*

$$T^{-2}\sum_{t=1}^{T} X_t^{(m)} X_t^{(m)*} \xrightarrow{w} \delta_m^2 C_m \int_0^1 W_m W_m^* du C_m^*, \tag{23}$$

$$T^{-2}\sum_{t=1}^{T} X_t^{(m)} X_t^{(n)*} \xrightarrow{P} 0, \quad z_n \neq z_m, \tag{24}$$

$$T^{-1}\sum_{t=1}^{T} X_t^{(m)} \varepsilon_t' \xrightarrow{w} \delta_m^2 C_m \int_0^1 W_m (dW_m)^*. \tag{25}$$

■

Next we want to find the asymptotic properties of the product moment matrices S_{00}, S_{10}, S_{11} and $S_{1\varepsilon}$. These are defined in terms of the residuals $U_{0t}, U_{1t}, U_{\varepsilon t}$ which in turn are defined in terms of the processes $X_{Rt}^{(1)}$, $X_{It}^{(1)}$ and $p(L)X_t$ corrected for $X_t^{(m)}$ for $m \neq (1,2)$.

In the following we let $z_1 = e^{i\theta}$, $0 < \theta < \pi$, and define $S_t^{(1)} = S_{Rt}^{(1)} + iS_{It}^{(1)}$, $X_t^{(1)} = X_{Rt}^{(1)} + iX_{It}^{(1)}$ and $C_1 = C_R^{(1)} + iC_I^{(1)}$. With the matrix representation of the complex processes, see Appendix A, we use the notation

$$\mathbf{X}_t^{(1)} = \begin{pmatrix} X_{Rt}^{(1)} & -X_{It}^{(1)} \\ X_{It}^{(1)} & X_{Rt}^{(1)} \end{pmatrix}, \mathbf{C}_1 = \begin{pmatrix} C_R^{(1)} & -C_I^{(1)} \\ C_I^{(1)} & C_R^{(1)} \end{pmatrix}, \mathbf{S}_t^{(1)} = \begin{pmatrix} S_{Rt}^{(1)} & -S_{It}^{(1)} \\ S_{It}^{(1)} & S_{Rt}^{(1)} \end{pmatrix}, \tag{26}$$

such that the complex representation

$$X_t^{(1)} = C_1 S_{t-1}^{(1)} \bar{z}_1^t + o_P(T^{\frac{1}{2}}),$$

in matrix notation becomes

$$\mathbf{X}_t^{(1)} = \begin{pmatrix} C_R^{(1)} & -C_I^{(1)} \\ C_I^{(1)} & C_R^{(1)} \end{pmatrix} \begin{pmatrix} S_{Rt-1}^{(1)} & -S_{It-1}^{(1)} \\ S_{It-1}^{(1)} & S_{Rt-1}^{(1)} \end{pmatrix} \begin{pmatrix} \cos(t\theta) & \sin(t\theta) \\ -\sin(t\theta) & \cos(t\theta) \end{pmatrix} + o_P(T^{\frac{1}{2}}).$$

From this we find by multiplying from the right by $(1,0)'$ that

$$\begin{pmatrix} X_{Rt}^{(1)} \\ X_{It}^{(1)} \end{pmatrix} = \mathbf{C}_1\mathbf{S}_{t-1}^{(1)} \begin{pmatrix} \cos(t\theta) \\ -\sin(t\theta) \end{pmatrix} + o_P(T^{\frac{1}{2}}).$$

Define the $\sigma-$field $\mathcal{F}_t$ as

$$\mathcal{F}_t = \sigma\left\{p(L)X_{t-1},\ldots,p(L)X_{t-k},\beta_m^* X_t^{(m)},\, m \neq (1,2)\right\},$$

that is, the σ-field generated by the stationary processes in the model equation except those that are derived from $X_{Rt}^{(1)}$ and $X_{It}^{(1)}$. Note that $\mathcal{F}_t$ is generated by variables before time t, since $X_t^{(m)}$ depends on lagged X_t.

We define the variances and covariances

$$Var\left\{\begin{array}{c} p(L)X_t \\ \boldsymbol{\beta}'\begin{pmatrix} X_{Rt}^{(1)} \\ X_{It}^{(1)} \end{pmatrix} \end{array}\middle| \mathcal{F}_t\right\} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0\beta} \\ \Sigma_{\beta 0} & \Sigma_{\beta\beta} \end{bmatrix}$$

Lemma 8 *The following identities hold*

$$\Sigma_{0\beta} = \tilde{\alpha}\Sigma_{\beta\beta}, \tag{27}$$

$$\Sigma_{00} = \tilde{\alpha}'\Sigma_{\beta\beta}\tilde{\alpha} + \Omega, \tag{28}$$

$$\Sigma_{\beta\beta.0}^{-1}\Sigma_{\beta 0}\Sigma_{00}^{-1} = \tilde{\alpha}'\Omega^{-1}, \tag{29}$$

$$\Sigma_{\beta\beta}^{-1} - \Sigma_{\beta\beta.0}^{-1} + \Sigma_{\beta\beta.0}^{-1}\Sigma_{\beta 0}\Sigma_{00}^{-1}\tilde{\alpha} = 0. \tag{30}$$

■

Theorem 9 *The asymptotic properties of the product moment matrices defined from $X_t^{(1)}$, corrected for the processes $X_t^{(m)}$, $m \neq 1,2$, are given by*

$$T^{-1}S_{11} \xrightarrow{w} \frac{1}{4}\mathbf{C}_1\int \mathbf{W}_1\mathbf{W}_1' du\mathbf{C}_1',\ 0 < \theta_1 < \pi,$$

$$S_{1\varepsilon} \xrightarrow{w} \frac{1}{2}\mathbf{C}_1\int \mathbf{W}_1(d\mathbf{W}_1)'\begin{pmatrix} I \\ 0 \end{pmatrix},\ 0 < \theta_1 < \pi,$$

where

$$\mathbf{W}_1 = \begin{pmatrix} W_R^{(1)} & -W_I^{(1)} \\ W_I^{(1)} & W_R^{(1)} \end{pmatrix}.$$

Furthermore we have the relations

$$\hat{\beta}'S_{11}\hat{\beta} \xrightarrow{P} \Sigma_{\beta\beta},\ \hat{\beta}'S_{10} \xrightarrow{P} \Sigma_{\beta 0},\ S_{00} \xrightarrow{P} \Sigma_{00},\ \hat{\beta}'S_{1\varepsilon} \xrightarrow{P} 0.$$

Note that S_{11} is $2p \times 2p$ but does not have complex structure. The limit of $T^{-1}S_{11}$, however, has complex structure as do the matrices $\mathbf{C}_1$ and $\mathbf{W}_1$. Note also that the $\mathbf{W}_1$ process appearing in Theorem 9 above is just the complex valued Brownian motion W_1 from Lemma 5 in the matrix representation of the complex process.

5 Asymptotic inference on rank and cointegrating relations

The main result about the estimator β is that it is asymptotically mixed Gaussian such that asymptotic inference on the coefficients can be conducted in the χ^2 distribution. The test statistic for hypotheses on the rank at seasonal frequency has a limit distribution, which is similar to the usual one, when expressed in terms of the complex Brownian motion

5.1 The asymptotic distribution of $\hat{\boldsymbol{\beta}}$

Although resorting to numerical algorithms for calculating $\hat{\boldsymbol{\beta}}$ is necessary, we can use the derived expression for the likelihood function (16) to obtain the asymptotic distribution of the maximum likelihood estimator. We do this by exploiting the fact that $\hat{\boldsymbol{\beta}}$ must be a solution to a set of first order conditions for maximizing (16).

The parameter $\boldsymbol{\beta}$ is not identified unless normalized in some way. This normalization can be accomplished by defining $\boldsymbol{\beta}_b = \boldsymbol{\beta}(\mathbf{b}'\boldsymbol{\beta})^{-1}$ for some $\mathbf{b}$ ($2p \times 2r$) of complex structure with the property that $\boldsymbol{\beta}'\mathbf{b}$ has full rank. For the analysis in the following it is convenient first to normalize the estimator on the true value $\boldsymbol{\beta}$ and choose $\mathbf{b} = \bar{\boldsymbol{\beta}} = \boldsymbol{\beta}(\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}$. We thus define $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\bar{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}})^{-1}$ and note that

$$\boldsymbol{\beta}'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = 0.$$

Thus we only have to investigate the limit of $T\bar{\boldsymbol{\beta}}'_{\perp}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$. We give the results for the model without deterministic terms and later mention how they are modified for the models (11) and (12).

Theorem 10 *The asymptotic distribution of the cointegration vector $\hat{\boldsymbol{\beta}}$ in the model with no deterministic terms is mixed Gaussian and given by*

$$T\bar{\boldsymbol{\beta}}'_{\perp}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{w} [\int_0^1 \mathbf{F}\mathbf{F}'du]^{-1} \int_0^1 \mathbf{F}(d\mathbf{V})',$$

where

$$\mathbf{F} = \boldsymbol{\beta}_\perp' \mathbf{C}_1 \mathbf{W}_1$$

and

$$\mathbf{V} = (\boldsymbol{\alpha}'\boldsymbol{\Omega}_\mathbf{c}^{-1}\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}'\boldsymbol{\Omega}_\mathbf{c}^{-1}\mathbf{W}_1.$$

where

$$\boldsymbol{\Omega}_c = \begin{pmatrix} \Omega & 0 \\ 0 & \Omega \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} \alpha_R & -\alpha_I \\ \alpha_I & \alpha_R \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_R & -\beta_I \\ \beta_I & \beta_R \end{pmatrix}.$$

Thus there is some redundancy built into the result, but the notation is chosen such that it coincides with the usual one for the unit root $z = 1$, see Johansen (1991).

Proof. The proof that the maximum likelihood estimator is consistent can be given along the same lines as the proof of consistency in Johansen (1996), where it is pointed out that due to the fact that the cointegration model is a sub-model of a Gaussian regression model, it is possible to find an upper bound of the likelihood function outside a neighborhood of the true value. This can then be applied to prove consistency. In the following we assume that $\tilde{\boldsymbol{\beta}}$ exists and is consistent.

The concentrated likelihood function is given by

$$-2\log L(\boldsymbol{\beta}) = T\log\frac{|\boldsymbol{\beta}'S_{11.0}\boldsymbol{\beta}|}{|\boldsymbol{\beta}'S_{11}\boldsymbol{\beta}|} + T\log|S_{00}|.$$

We next want to make an expansion of the likelihood function around the maximum, and we use the expansion

$$\log|(x+h)'A(x+h)| = \log|x'Ax| + 2tr\{(x'Ax)^{-1}x'Ah\} + O(|h|^2). \tag{31}$$

This gives the first order condition

$$tr\{[(\tilde{\boldsymbol{\beta}}'S_{11}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11} - (\tilde{\boldsymbol{\beta}}'S_{11.0}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11.0}]\mathbf{h}\} = 0,$$

for all $\mathbf{h}$ of complex structure. This implies that

$$[(\tilde{\boldsymbol{\beta}}'S_{11}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11} - (\tilde{\boldsymbol{\beta}}'S_{11.0}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11.0}]^c = 0, \tag{32}$$

where $[\ldots]^c$ denotes the complexified matrix, see Appendix A. We first find the weak limit for the matrix in (32) before it is complexified. Multiplying from the right by $\boldsymbol{\beta}_\perp$, which has complex structure, we find

$$\begin{aligned}
&(\tilde{\boldsymbol{\beta}}'S_{11}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11}\boldsymbol{\beta}_\perp - (\tilde{\boldsymbol{\beta}}'S_{11.0}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{11.0}\boldsymbol{\beta}_\perp \\
&= ((\tilde{\boldsymbol{\beta}}'S_{11}\tilde{\boldsymbol{\beta}})^{-1} - (\tilde{\boldsymbol{\beta}}'S_{11.0}\tilde{\boldsymbol{\beta}})^{-1})\tilde{\boldsymbol{\beta}}'S_{11}\boldsymbol{\beta}_\perp + (\tilde{\boldsymbol{\beta}}'S_{11.0}\tilde{\boldsymbol{\beta}})^{-1}\tilde{\boldsymbol{\beta}}'S_{10}S_{00}^{-1}S_{01}\boldsymbol{\beta}_\perp \\
&= (\Sigma_{\beta\beta}^{-1} - \Sigma_{\beta\beta.0}^{-1})\tilde{\boldsymbol{\beta}}'S_{11}\boldsymbol{\beta}_\perp + \Sigma_{\beta\beta.0}^{-1}\Sigma_{\beta0}\Sigma_{00}^{-1}S_{01}\boldsymbol{\beta}_\perp + o_P(1).
\end{aligned}$$

We have here applied the results in Theorem 9 that $\tilde{\boldsymbol{\beta}}' S_{11}\tilde{\boldsymbol{\beta}} \xrightarrow{P} \Sigma_{\beta\beta}$, $\tilde{\boldsymbol{\beta}}' S_{11.0}\tilde{\boldsymbol{\beta}} \xrightarrow{P} \Sigma_{\beta 0}$ and $S_{00} \xrightarrow{P} \Sigma_{00}$. From (15) we find

$$S_{01}\boldsymbol{\beta}_\perp = \tilde{\alpha}\boldsymbol{\beta}' S_{11}\boldsymbol{\beta}_\perp + S_{\varepsilon 1}\boldsymbol{\beta}_\perp = \tilde{\alpha}(\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}})' S_{11}\boldsymbol{\beta}_\perp + \tilde{\alpha}\tilde{\boldsymbol{\beta}}' S_{11}\boldsymbol{\beta}_\perp + S_{\varepsilon 1}\boldsymbol{\beta}_\perp.$$

Inserting this above we find

$$\begin{aligned}
&(\Sigma_{\beta\beta}^{-1} - \Sigma_{\beta\beta.0}^{-1} + \Sigma_{\beta\beta.0}^{-1}\Sigma_{\beta 0}\Sigma_{00}^{-1}\tilde{\alpha})\tilde{\boldsymbol{\beta}}' S_{11}\boldsymbol{\beta}_\perp \\
&+\Sigma_{\beta\beta.0}^{-1}\Sigma_{\beta 0}\Sigma_{00}^{-1}(\tilde{\alpha}(\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}})'\bar{\boldsymbol{\beta}}_\perp\boldsymbol{\beta}_\perp' S_{11}\boldsymbol{\beta}_\perp + S_{\varepsilon 1}\boldsymbol{\beta}_\perp) \\
&= -\tilde{\alpha}'\Omega^{-1}(\tilde{\alpha}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'\bar{\boldsymbol{\beta}}_\perp\boldsymbol{\beta}_\perp' S_{11}\boldsymbol{\beta}_\perp + S_{\varepsilon 1}\boldsymbol{\beta}_\perp),
\end{aligned}$$

since the first term is zero by (30) and the coefficient simplifies by (29). The weak limit of this is

$$-\tilde{\alpha}'\Omega^{-1}\left[\tilde{\alpha}\mathbf{B}'\frac{1}{4}\int_0^1 \mathbf{F}\mathbf{F}' du + \frac{1}{2}(I,0)\int_0^1 (d\mathbf{W}_1)\mathbf{F}'\right],$$

where we have used the notation $\mathbf{B}$ for the weak limit of $\mathbf{B}_T = T\bar{\boldsymbol{\beta}}_\perp'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$, see (29). Thus the limit of (32) becomes

$$\left[\tilde{\alpha}'\Omega^{-1}\tilde{\alpha}\mathbf{B}'\frac{1}{4}\int_0^1 \mathbf{F}\mathbf{F}' du - \frac{1}{2}\tilde{\alpha}'\Omega^{-1}(I,0)\int_0^1 (d\mathbf{W}_1)\mathbf{F}'\right]^c = 0.$$

We still have to simplify this result before we can solve the equation for $\mathbf{B}$. We get, since $\mathbf{B}'\int_0^1 \mathbf{F}\mathbf{F}' du$ and $\int_0^1 (d\mathbf{W}_1)\mathbf{F}'$ have complex structure, that the first order condition (32) is equivalent to

$$\left[\tilde{\alpha}'\Omega^{-1}\tilde{\alpha}\right]^c \mathbf{B}'\frac{1}{4}\int_0^1 \mathbf{F}\mathbf{F}' du - \frac{1}{2}\left[\tilde{\alpha}'\Omega^{-1}(I,0)\right]^c\int_0^1 (d\mathbf{W}_1)\mathbf{F}' = 0.$$

This is now solved for $\mathbf{B}$ and we therefore want to find the two complexified matrices. We find

$$[\tilde{\alpha}'\Omega^{-1}\tilde{\alpha}]^c = 2\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha}$$

and

$$\left[\tilde{\alpha}'\Omega^{-1}(I,0)\right]^c = \boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1},$$

for

$$\boldsymbol{\Omega}_c = \begin{pmatrix} \Omega & 0 \\ 0 & \Omega \end{pmatrix},$$

which shows that

$$\mathbf{B} = \left(\int_0^1 \mathbf{F}\mathbf{F}' du\right)^{-1}\int_0^1 \mathbf{F}(d\mathbf{W}_1)'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha}(\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha})^{-1}.$$

■

Next we give a result for the estimator of $\boldsymbol{\beta}$ normalized on a matrix $\mathbf{b}$, that is, $\boldsymbol{\beta}_b = \boldsymbol{\beta}(\mathbf{b}'\boldsymbol{\beta})^{-1}$.

Theorem 11 *Consider the seasonal frequency $z_1 = e^{i\theta_1}$, and the matrix $\boldsymbol{\beta}$ with complex structure normalized by $\boldsymbol{\beta}'\mathbf{b} = \mathbf{I}$. In the model with no deterministic terms $\hat{\boldsymbol{\beta}}_b$ is consistent and asymptotically mixed Gaussian,*

$$T(\hat{\boldsymbol{\beta}}_b - \boldsymbol{\beta}) \xrightarrow{w} (I - \boldsymbol{\beta}\mathbf{b}')\boldsymbol{\beta}_\perp[\int_0^1 \mathbf{F}_1\mathbf{F}_1' du]^{-1} \int \mathbf{F}_1(d\mathbf{V}_1)', \tag{33}$$

where

$$\begin{aligned} \mathbf{F}_1(u) &= \boldsymbol{\beta}_\perp'\mathbf{C}_1\mathbf{W}_1(u), \\ \mathbf{V}_1(u) &= (\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\mathbf{W}_1(u). \end{aligned}$$

The asymptotic conditional variance matrix is

$$(I - \boldsymbol{\beta}\mathbf{b}')\boldsymbol{\beta}_\perp[\int_0^1 \mathbf{F}_1\mathbf{F}_1' du]^{-1}\boldsymbol{\beta}_\perp'(\mathbf{I} - \mathbf{b}\boldsymbol{\beta}') \otimes (\boldsymbol{\alpha}'\boldsymbol{\Omega}_\mathbf{c}^{-1}\boldsymbol{\alpha})^{-1}, \tag{34}$$

which by Theorem 9 is estimated consistently by

$$T(I - \hat{\boldsymbol{\beta}}_b\mathbf{b}')\hat{\boldsymbol{\beta}}_\perp[4\hat{\boldsymbol{\beta}}_\perp' S_{11}\hat{\boldsymbol{\beta}}_\perp]^{-1}\hat{\boldsymbol{\beta}}_\perp'(I - \mathbf{b}\hat{\boldsymbol{\beta}}_b') \otimes (\hat{\boldsymbol{\alpha}}_b'\hat{\boldsymbol{\Omega}}_c^{-1}\hat{\boldsymbol{\alpha}}_b)^{-1}. \tag{35}$$

Thus linear and non-linear hypotheses on the coefficients of the just identified vector $\boldsymbol{\beta}_b$ can be tested asymptotically by construction of $t-$ ratios using (35) as variance matrix.

Proof. The proof of (33) follows from Theorem 10 by the expansion

$$\hat{\boldsymbol{\beta}}_b = (I - \boldsymbol{\beta}(\mathbf{b}'\boldsymbol{\beta})^{-1}\mathbf{b}')(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\mathbf{b}'\boldsymbol{\beta})^{-1} + O_P(|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}|^2).$$

The proof that (34) is a consistent estimator follows from Theorem (9). ■

If we instead consider the models (11) or (12) we get much the same results. A detailed study will show that the estimated cointegrating vectors $\hat{\beta}_m$ are T consistent but their extension $\hat{\rho}_m$ is only $T^{\frac{1}{2}}$ consistent. This gives some difficulties in the formulation, but the end result is that one can treat the full extended vector as asymptotically Gaussian with a variance matrix given by (35), see Harbo *et al.* (1996) for the details in the case of zero frequency.

5.2 Test for cointegrating rank

This section contains a test to determine the rank r of β at the seasonal frequency $z_1 = e^{i\theta}$. We here concentrate on deriving the result for testing at strictly complex frequencies, which yields a result similar to the usual test but involving complex Brownian motions. We focus on the model without deterministic terms and give the results for the other cases without proof.

Theorem 12 *The asymptotic distribution of the test statistic for the hypothesis of $r < p$ cointegrating relations at complex seasonal frequency is asymptotically distributed as*

$$\frac{1}{2}tr\{\int_0^1 (d\mathbf{B})\mathbf{B}'(\int_0^1 \mathbf{B}\mathbf{B}'du)^{-1}\int_0^1 \mathbf{B}(d\mathbf{B}')\}, \tag{36}$$

where $\mathbf{B}$ *is standard complex Brownian motion of dimension* $2(p-r)$

$$\mathbf{B} = \begin{pmatrix} B_R & -B_I \\ B_I & B_R \end{pmatrix}.$$

The distribution is tabulated by simulation in Table 1

Proof. From (16) we find that the maximized likelihood function for $p = r$ becomes

$$L_{\max}^{-\frac{2}{T}} = |S_{00}|\frac{|S_{11.0}|}{|S_{11}|}, \tag{37}$$

and hence that the likelihood ratio test statistics can be found as

$$Q^{-\frac{2}{T}}(H(r)|H(p)) = \frac{|S_{11}||\hat{\boldsymbol{\beta}}'S_{11.0}\hat{\boldsymbol{\beta}}|}{|S_{11.0}||\hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}|}.$$

Now choose $\hat{\boldsymbol{\beta}}_\perp$ orthogonal to $\hat{\boldsymbol{\beta}}$ and use the identities

$$\begin{aligned}
&|(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}}_\perp)'||S_{11}||(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}}_\perp)'| \\
&= |(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}}_\perp)'S_{11}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\beta}}_\perp)'| \\
&= \begin{vmatrix} \hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}} & \hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}_\perp \\ \hat{\boldsymbol{\beta}}_\perp'S_{11}\hat{\boldsymbol{\beta}} & \hat{\boldsymbol{\beta}}_\perp'S_{11}\hat{\boldsymbol{\beta}}_\perp \end{vmatrix} \\
&= \left|\hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}\right|\left|\hat{\boldsymbol{\beta}}_\perp'S_{11}\hat{\boldsymbol{\beta}}_\perp - \hat{\boldsymbol{\beta}}_\perp'S_{11}\hat{\boldsymbol{\beta}}\left(\hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}\right)^{-1}\hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}_\perp\right| \\
&= \left|\hat{\boldsymbol{\beta}}'S_{11}\hat{\boldsymbol{\beta}}\right|\left|\hat{\boldsymbol{\beta}}_\perp'S_{11.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp\right|
\end{aligned}$$

and a similar one for the matrix $S_{11.0}$ to prove the expression

$$-2\log Q(H(r)|H(p)) = -T\log\frac{\left|\hat{\boldsymbol{\beta}}_\perp'S_{11.0\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp\right|}{\left|\hat{\boldsymbol{\beta}}_\perp'S_{11.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp\right|}. \tag{38}$$

The idea of the proof is to derive the asymptotic distribution of (38) by noting that it is a function of $\hat{\beta}$, for which the distribution is derived in Theorem 11.

From the consistency of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_\perp$ it follows from Theorem 9 that

$$T^{-1}\hat{\boldsymbol{\beta}}_\perp' S_{11.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp \xrightarrow{w} \frac{1}{4}\boldsymbol{\beta}_\perp'\mathbf{C}_1\int_0^1 \mathbf{W}_1\mathbf{W}_1' du \mathbf{C}_1'\boldsymbol{\beta}_\perp = \frac{1}{4}\int_0^1 \mathbf{F}\mathbf{F}' du,$$

and the same result holds for $\hat{\boldsymbol{\beta}}_\perp' S_{11.0\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp$. Thus the ratio in (38) tends to 1, and we therefore get the result, applying (31),

$$\begin{aligned}
&-2\log Q(H(r)|H(p)) \\
&= -T\log|I - (\hat{\boldsymbol{\beta}}_\perp' S_{11.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp)^{-1}\hat{\boldsymbol{\beta}}_\perp' S_{10.\hat{\beta}}S_{00.\hat{\beta}}^{-1}S_{01.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp| \\
&= tr\left\{(T^{-1}\hat{\boldsymbol{\beta}}_\perp' S_{11.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp)^{-1}\hat{\boldsymbol{\beta}}_\perp' S_{10.\hat{\beta}}S_{00.\hat{\beta}}^{-1}S_{01.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp\right\} + o_P(1).
\end{aligned}$$

We want to find the limit of this quantity. First we consider

$$\begin{aligned}
&S_{00.\hat{\beta}} = S_{00} - S_{01}\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}' S_{11}\hat{\boldsymbol{\beta}})^{-1}\hat{\boldsymbol{\beta}}' S_{10} \\
&\xrightarrow{P} \Sigma_{00} - \Sigma_{0\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta 0} = \Omega + \tilde{\alpha}\Sigma_{\beta\beta}\tilde{\alpha}' - \Sigma_{0\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta 0} = \Omega.
\end{aligned}$$

Next we consider

$$\begin{aligned}
&\hat{\boldsymbol{\beta}}_\perp' S_{10.\hat{\beta}} = \hat{\boldsymbol{\beta}}_\perp' S_{10} - \hat{\boldsymbol{\beta}}_\perp' S_{11}\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}' S_{11}\hat{\boldsymbol{\beta}})^{-1}\hat{\boldsymbol{\beta}}' S_{10} \\
&= \hat{\boldsymbol{\beta}}_\perp' S_{1\varepsilon} + \hat{\boldsymbol{\beta}}_\perp' S_{11}\boldsymbol{\beta}\tilde{\alpha}' - \hat{\boldsymbol{\beta}}_\perp' S_{11}\hat{\boldsymbol{\beta}}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta 0} + o_P(1) \\
&= \boldsymbol{\beta}_\perp' S_{1\varepsilon} - \boldsymbol{\beta}_\perp' S_{11}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\tilde{\alpha}' + o_P(1) \\
&= \boldsymbol{\beta}_\perp' S_{1\varepsilon} - \boldsymbol{\beta}_\perp' S_{11}\boldsymbol{\beta}_\perp\bar{\boldsymbol{\beta}}_\perp'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\tilde{\alpha}' + o_P(1)
\end{aligned}$$

From Theorem 9 we find that this converges towards

$$\begin{aligned}
&\tfrac{1}{2}\textstyle\int_0^1 \mathbf{F}(d\mathbf{W}_1')\begin{pmatrix} I \\ 0 \end{pmatrix} - \tfrac{1}{4}\int_0^1 \mathbf{F}(d\mathbf{W}_1')\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha}(\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha})^{-1}\tilde{\alpha}' \\
&= \tfrac{1}{2}\textstyle\int_0^1 \mathbf{F}(d\mathbf{W}_1')(I - \boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha}(\boldsymbol{\alpha}'\boldsymbol{\Omega}_c^{-1}\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}')\begin{pmatrix} I \\ 0 \end{pmatrix} \\
&= \tfrac{1}{2}\textstyle\int_0^1 \mathbf{F}(d\mathbf{W}_1')\boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\begin{pmatrix} I \\ 0 \end{pmatrix}.
\end{aligned}$$

Thus we find that

$$\hat{\boldsymbol{\beta}}_\perp' S_{10.\hat{\beta}}S_{00.\hat{\beta}}S_{01.\hat{\beta}}\hat{\boldsymbol{\beta}}_\perp \xrightarrow{w} \frac{1}{2}\int_0^1 \mathbf{F}(d\mathbf{W}_1')M\int_0^1 (d\mathbf{W}_1')\mathbf{F}\frac{1}{2},$$

where M is given by

$$M = \boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\begin{pmatrix} I \\ 0 \end{pmatrix}\Omega^{-1}\begin{pmatrix} I \\ 0 \end{pmatrix}\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp',$$

such that

$$
\begin{aligned}
M^c &= \frac{1}{2}\boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\Omega}_c^{-1}\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp' \\
&= \frac{1}{2}\boldsymbol{\alpha}_\perp(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_c\boldsymbol{\alpha}_\perp)^{-1}\boldsymbol{\alpha}_\perp',
\end{aligned}
$$

since

$$
\left[\begin{pmatrix} I \\ 0 \end{pmatrix}\Omega^{-1}\begin{pmatrix} I \\ 0 \end{pmatrix}\right]^c = \frac{1}{2}\boldsymbol{\Omega}_c^{-1}.
$$

The asymptotic distribution is then given by

$$
\begin{aligned}
& tr\{\left[\int_0^1 \mathbf{FF}'du\right]^{-1}\int_0^1 \mathbf{F}(\mathbf{dW}_1')M\int_0^1(d\mathbf{W}_1)\mathbf{F}'\} \\
= \;& tr\{\left[\int_0^1 \mathbf{FF'du}\right]^{-1}\int_0^1 \mathbf{F}(\mathbf{dW}_1')M^c\int_0^1(\mathbf{dW}_1)\mathbf{F}'\},
\end{aligned}
$$

since both the matrices $\int_0^1 \mathbf{FF}'du$ and $\int_0^1 \mathbf{F}(\mathbf{dW}_1')$ have complex structure. Combining the results we find that

$$
-2\log Q(H(r)|H(p)) \xrightarrow{w} \frac{1}{2}tr\{\int_0^1(d\mathbf{B})\mathbf{B}')\left[\int_0^1 \mathbf{BB'du}\right]^{-1}\int_0^1 \mathbf{B}(\mathbf{dB}')\},
$$

where

$$
\mathbf{B} = (\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}_\mathbf{c}\boldsymbol{\alpha}_\perp)^{-\frac{1}{2}}\boldsymbol{\alpha}_\perp'\mathbf{W}_1.
$$

■

By choosing to express the result in terms of the complex Brownian motion we find that, apart from the factor $\frac{1}{2}$, the result looks very much like the result for the real case, see Johansen (1991), for $z = 1$, and Lee (1992) for the case $z = -1$. The result given in (36) corresponds to formula (3.35) in Lee (1992). The calculations of the likelihood ratio statistics (3.34), however, are not correct and there is an error in the proof giving the asymptotic properties. The choice of δ_q cannot be made as stated just below (A.42). The resulting formula for the limit distribution is, however, correct.

Finally consider the test for cointegration rank at complex frequency when there are deterministic terms in the model.

Theorem 13 *In model (11) the asymptotic distribution of the test statistic for the hypothesis of* $r < p$ *cointegrating relations at complex seasonal frequency is asymptotically distributed as*

$$
\frac{1}{2}tr\{\int_0^1(d\mathbf{B})\mathbf{H}'(\int_0^1 \mathbf{HH}'du)^{-1}\int_0^1 \mathbf{H}(d\mathbf{B}')\}, \tag{39}
$$

where $\mathbf{B}$ *is standard complex Brownian motion of dimension* $2(p-r)$ *and* $\mathbf{H} = (\mathbf{B}', I)'$. *The limit distribution is tabulated in Table 2.*

Note how the properties of the extended process $(X_t^{(m)\prime}, z_m^t)'$ are reflected in the extended Brownian motion $\mathbf{H}$.

Finally if we consider model (12) which allows for a linear trend in the process we find the same result but with the definition of $\mathbf{H}$ changed.

Theorem 14 *In model (12) the asymptotic distribution of the test statistic for the hypothesis of* $r < p$ *cointegrating relations at complex seasonal frequency is asymptotically distributed as*

$$\frac{1}{2} tr\{\int_0^1 (d\mathbf{B})\mathbf{H}'(\int_0^1 \mathbf{H}\mathbf{H}'du)^{-1}\int_0^1 \mathbf{H}(d\mathbf{B}')\}, \tag{40}$$

where $\mathbf{B}$ *is standard complex Brownian motion of dimension* $2(p-r)$ *and* $\mathbf{H} = (\mathbf{B}' - \bar{\mathbf{B}}', I)'$.

Again the process $\mathbf{H}$ reflects the properties of the extended process $X_t^{(m)}$, but this time corrected for the average $\bar{\mathbf{B}}$ corresponding to fitting an unrestricted constant in the equations. The limit distribution is tabulated by simulation in Table 3.

6 References

Boswijk, H. P., 1995, Identifiability of cointegrated systems, Tinbergen Institute Discussion paper, TI, 95-78.

Chan, N. H. and C. Z. Wei, 1988, Limiting distributions of least squares estimators of unstable autoregressive processes, The Annals of Statistics 16, 367-401.

Dahl Pedersen, I.,1996, A practical implementation of the seasonal cointegration theory, 93 pp., Masters Thesis, University of Copenhagen.

Engle R. F., Granger, C. W. J., Hylleberg, S., and H. S. Lee, 1993, Seasonal cointegration: The Japanese consumption function, Journal of Econometrics, 55, 275-298.

Gregoir, S., 1993*a*, Multivariate time series with various hidden unit roots: Part II: Estimation and Test, Discussion paper.

Gregoir, S., 1993*b*, Multivariate time series with various hidden unit roots: Part I. Integral operator algebra and representation theorem, Discussion paper.

Harbo, I., Johansen S., Nielsen, B. and A. Rahbek, 1996, Asymptotic inference on cointegrating rank in partial systems, Discussion paper, University of Copenhagen.

Hylleberg, S., Engle R. F., Granger, C. W. J. and S. B., Yoo, 1990, Seasonal integration and cointegration, Journal of Econometrics, 44, 215-238.

Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models Econometrica, 59, 1551-1580.

Johansen, S., 1996, Likelihood analysis of the $I(2)$ model, forthcoming Scandinavian Journal of Statistics.

Kunst R. M. and , P. H. Franses, 1995, On the role of seasonal intercepts in seasonal cointegration, IHS working paper Reihe Ökonomie no. 15.

Lee, S. L., 1992, Maximum likelihood inference on cointegration and seasonal cointegration, Journal of Econometrics, 54, 1-47.

Schaumburg, E., 1996, Inference in seasonally integrated VAR models, 98 pp. Masters thesis, University of Copenhagen.

Appendix A

A 1. Complex matrices and real matrices with complex structure

Complex number $z = a + ib$ can be represented by the matrix

$$\mathbf{z} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

in the sense that this representation preserves linear operations and also complex multiplication, that is, if

$$(a + ib)(c + id) = e + if,$$

then

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} e & -f \\ f & e \end{pmatrix}.$$

We represent the complex $p \times q$ matrix $= A + iB$ by the real $2p \times 2q$ matrix $\mathbf{F}$

$$\mathbf{F} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.$$

Throughout we use boldface to denote the real matrices with this complex structure. Note that if $F^* = A' - iB'$ then F^* has the representation

$$\mathbf{F}' = \begin{pmatrix} A' & B' \\ -B' & A' \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}'.$$

We shall say that F is complex, but that $\mathbf{F}$ has complex structure.

We consider the transformation of a $2p \times 2q$ matrix to a matrix of complex structure given by

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &\rightarrow \frac{1}{2}\left[\begin{pmatrix} A & B \\ C & D \end{pmatrix} + \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\right] \\ &= \frac{1}{2}\begin{pmatrix} A+D & B-C \\ C-B & A+D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{\mathbf{c}}. \end{aligned}$$

We can discuss this by the transformation

$$\mathbf{I} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

such that $\mathbf{I}^2 = -\mathbf{I}$ and $\imath' = -\mathbf{I}$ Then for a $2p \times 2q$ matrix M we have

$$M^c = \frac{1}{2}\left(M + \mathbf{I}M\mathbf{I}'\right).$$

and if $\mathbf{M}$ has complex structure then $\mathbf{IMI}' = \mathbf{M}$, such that $\mathbf{M} = \mathbf{M}^c$. If M is any $2p \times 2q$ matrix and $\mathbf{h}$ has complex structure then

$$(M\mathbf{h})^c = \frac{1}{2}(M\mathbf{h} + \mathbf{I}M\mathbf{h}\mathbf{I}') = \frac{1}{2}(M\mathbf{h} + \mathbf{I}M\mathbf{I}'\mathbf{I}\mathbf{h}\mathbf{I}') = \frac{1}{2}(M + \mathbf{I}M\mathbf{I}')\mathbf{h} = M^c\mathbf{h}.$$

Finally notice that if $tr\{M\mathbf{h}\} = 0$ for all $\mathbf{h}$ with complex structure then $M^c = 0$, since

$$tr\{M\mathbf{h}\} = tr\{(M\mathbf{h})^c\} = tr\{M^c\mathbf{h}\} = 0$$

for all $\mathbf{h}$ with complex structure implies that $M^c = 0$.

A 2. Asymptotics

This appendix contains brief proofs of some of the technical results stated in Section 4.

Proof of Theorem 6. The first result (20) follows by the continuous mapping theorem and the second (21) by noting that $\Delta\bar{S}_t^{(n)} = \bar{z}_n^t \varepsilon_t$.

The third result (22) follows by a partial summation. Let $|A|^2 = tr\{A^* A\}$ for a complex matrix, and let $c = \sup_t |F(t)|$. Then

$$\begin{aligned}
&T^{-2}\textstyle\sum_{t=1}^T S_t^{(m)} f(t) S_t^{(n)*} \\
&= T^{-2}\textstyle\sum_{t=1}^T S_t^{(m)}(F(t) - F(t-1))S_t^{(n)*} \\
&= T^{-2}\textstyle\sum_{t=1}^T S_t^{(m)} F(t) S_t^{(n)*} - T^{-2}\sum_{t=1}^T (S_{t-1}^{(m)} + \Delta S_t^{(m)})F(t-1)(S_{t-1}^{(n)} + \Delta S_t^{(n)})^* \\
&= T^{-2} S_T^{(m)} F(T) S_T^{(n)*} - T^{-2}\textstyle\sum_{t=1}^T S_{t-1}^{(m)} F(t-1)\Delta S_t^{(n)*} \\
&-T^{-2}\textstyle\sum_{t=1}^T \Delta S_t^{(m)} F(t-1) S_{t-1}^{(n)*} - T^{-2}\sum_{t=1}^T \Delta S_t^{(m)} F(t-1)\Delta S_t^{(n)*}.
\end{aligned} \tag{41}$$

The first term is written as

$$(T^{-\frac{1}{2}} S_T^{(m)}) T^{-1} F(T) (T^{-\frac{1}{2}} S_T^{(n)*}) \xrightarrow{P} 0,$$

since F is bounded and $T^{-\frac{1}{2}} S_T^{(m)}$ converges weakly. The second and third terms are evaluated as follows:

$$\begin{aligned}
&E|T^{-2}\textstyle\sum_{t=1}^T S_{t-1}^{(m)} F(t-1)\Delta S_t^{(n)*}| \\
&\leq cT^{-2}\textstyle\sum_{t=1}^T E|S_{t-1}^{(m)}| E|\Delta S_t^{(n)*}| \\
&\leq c_1 T^{-2}\textstyle\sum_{t=1}^T t^{\frac{1}{2}} \in O(T^{-\frac{1}{2}}).
\end{aligned}$$

Thus the second and third term tend to zero, and the last term is evaluated as

$$\begin{aligned}
&T^{-2} E|\textstyle\sum_{t=1}^T \Delta S_t^{(m)} F(t-1)\Delta S_t^{(n)*}| \\
&\leq cT^{-2}\textstyle\sum_{t=1}^T E|\Delta S_t^{(m)}| E|\Delta S_t^{(n)*}| \in O(T^{-1}).
\end{aligned}$$

■

Proof of Corollary 7. The relation

$$T^{-2}\sum_{t=1}^{T} X_t^{(m)} X_t^{(n)*} = T^{-2}\sum_{t=1}^{T} \bar{z}_m^t z_n^t C_m S_{t-1}^{(m)} S_{t-1}^{(n)*} C_n^* + o_P(1),$$

shows that the asymptotic behaviour of the product moments depends on the boundedness of

$$\sum_{t=0}^{T} \bar{z}_m^t z_n^t = \frac{1-(\bar{z}_m z_n)^{T+1}}{1-\bar{z}_m z_n},$$

which remains bounded by if $z_m \neq z_n$. Thus for $z_m \neq z_n$ the product moment will converge to zero, whereas for $z_m = z_n$ we get the limit stated, which proves (23) and (24). The result (25) follows from (21). ∎

Thus the reason that the mixed moments tend to zero is not that they are asymptotically independent (which they are) but the factor $\bar{z}_m^t z_n^t$ which appears in the summation. The factor $\bar{z}_m^t$ comes from the representation of $X_t^{(m)}$ and also implies that the limit of $T^{-1}\sum_{t=1}^{T} X_t^{(m)}\varepsilon_t'$ does not involve the limit of $T^{-\frac{1}{2}}\sum_{t=1}^{T}\varepsilon_t$ but rather the limit of $T^{-\frac{1}{2}}\sum_{t=1}^{T}\bar{z}_m^t\varepsilon_t$.

Proof of Lemma 8. From the model equations

$$p(L)X_t = \tilde{\alpha}\beta' \begin{pmatrix} X_{Rt}^{(1)} \\ X_{It}^{(1)} \end{pmatrix} + \sum_{m=3}^{s} \alpha_m \beta_m^* X_t^{(m)} + \sum_{j=1}^{k} \Gamma_j p(L) X_{t-j} + \varepsilon_t,$$

it follows by taking conditional variances and covariances given the lagged values of $p(L)X_t$ and the remaining linear combinations $\beta_m^* X_t^{(m)}$ that (27) and (28) hold.

In order to prove (29) we write it as

$$\Sigma_{\beta 0} = \Sigma_{\beta\beta.0}\tilde{\alpha}'\Omega^{-1}\Sigma_{00},$$

and introduce the normalized vector

$$u = \Omega^{-\frac{1}{2}}\tilde{\alpha}\Sigma_{\beta\beta}^{\frac{1}{2}}.$$

After some reductions, applying $\Sigma_{\beta 0} = \tilde{\alpha}\Sigma_{\beta\beta}$ the relation (29) reduces to

$$u' = (I - u'(I+uu')^{-1}u)u'(I+uu'),$$

which follows from the identity

$$u'(I+uu')^{-1}u = (u'u)(I+u'u)^{-1}.$$

Next we multiply in (30) by $\Sigma_{\beta\beta.0}$ and $\Sigma_{\beta\beta}$ and find

$$\Sigma_{\beta\beta.0} - \Sigma_{\beta\beta} + \Sigma_{\beta 0}\Sigma_{00}^{-1}\Sigma_{0\beta} = 0,$$

which is zero by the definition of $\Sigma_{\beta\beta.0}$. ∎

Proof of Theorem 9. We first give a result for product moments of $X_t^{(m)}$ and $X_t^{(n)}$ corresponding to complex roots z_m and z_n :

$$T^{-1}M_{11}^{(n,m)} = T^{-2}\sum_{t=1}^{T}\begin{pmatrix} X_{Rt}^{(m)} \\ X_{It}^{(m)} \end{pmatrix}\begin{pmatrix} X_{Rt}^{(n)} \\ X_{It}^{(n)} \end{pmatrix}'$$

$$= T^{-2}\sum_{t=1}^{T}\mathbf{C}_m\mathbf{S}_{t-1}^{m}\begin{pmatrix} \cos(t\theta_m)\cos(t\theta_n) & -\cos(t\theta_m)\sin(t\theta_n) \\ -\sin(t\theta_m)\cos(t\theta_n) & \sin(t\theta_m)\sin(t\theta_n) \end{pmatrix}\mathbf{S}_{t-1}^{n\prime}\mathbf{C}_m' + o_P(1).$$

The matrix in the middle is

$$\frac{1}{2}\begin{pmatrix} \cos((\theta_m-\theta_n)t)+\cos((\theta_m+\theta_n)t) & \sin((\theta_m-\theta_n)t)-\sin((\theta_m+\theta_n)t) \\ \sin((\theta_n-\theta_m)t)-\sin((\theta_m+\theta_n)t) & \cos((\theta_m-\theta_n)t)-\cos((\theta_m+\theta_n)t)) \end{pmatrix},$$

which remain bounded when summed unless $\theta_m = \theta_n$, in which case the matrix equals

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \cos(2t\theta_m) & -\sin(2t\theta_m) \\ -\sin(2t\theta_m) & -\cos(2t\theta_m) \end{pmatrix},$$

where the last term is bounded when summed.

Hence $T^{-1}M_{11}^{(n,m)}$ converges to zero if $n \neq m$ and for $n = m$ has the same limit as

$$\frac{1}{2}T^{-2}\sum_{t=1}^{T}\mathbf{C}_m\mathbf{S}_{t-1}^{(m)}\mathbf{S}_{t-1}^{(m)\prime}\mathbf{C}_m'$$

$$\xrightarrow{w} \frac{1}{4}\mathbf{C}_m\int_0^1 \mathbf{W}_m\mathbf{W}_m' du\mathbf{C}_m'.$$

Similarly we find that for z_m complex

$$\begin{aligned} M_{1\varepsilon}^{(m)} &= T^{-1}\textstyle\sum_{t=1}^{T}\begin{pmatrix} X_{Rt}^{(m)} \\ X_{It}^{(m)} \end{pmatrix}\varepsilon_t' \\ &= \mathbf{C}_m(T^{-1}\textstyle\sum_{t=1}^{T}\mathbf{S}_{t-1}^{(m)}\begin{pmatrix} \cos(\theta_m t) \\ -\sin(\theta_m t) \end{pmatrix}\varepsilon_t') + o_P(1) \\ &= \mathbf{C}_m T^{-1}\textstyle\sum_{t=1}^{T}\mathbf{S}_{t-1}^{(m)}(\Delta\mathbf{S}_t^{(m)})'\begin{pmatrix} I \\ 0 \end{pmatrix} + o_P(1) \\ &\xrightarrow{w} \tfrac{1}{2}\mathbf{C}_m\int \mathbf{W}_m(d\mathbf{W}_m)'\begin{pmatrix} I \\ 0 \end{pmatrix}. \end{aligned}$$

If either z_m or z_n are real similar results can be proved. Finally we want the results for the product moment matrices constructed from the residuals U_t. It is clear that the limit of the product moments of $X_{Rt}^{(m)}$ and $X_{It}^{(m)}$ are not influenced by the preliminary regression on the lagged values of $p(L)X_t$, since these are stationary. The matrix S_{11} is $M_{11}^{(1,1)}$ corrected for the other processes. Since the mixed moments $T^{-1}M_{11}^{(n,m)}$converge to zero, the limit of $T^{-1}S_{11}$ is the same as that of $T^{-1}M_{11}^{(1,1)}$. Similarly the limit of $M_{1\varepsilon}^{(1)}$ is the same as that of $S_{1\varepsilon}$. ■

Appendix B

Tables

In this Appendix the asymptotic distributions of the likelihood ratio test statistics for cointegrating rank at complex frequency are tabulated. The limit distributions all have the form

$$\frac{1}{2}tr\left\{\int_0^1 (d\mathbf{B})\mathbf{H}'\left[\int_0^1 \mathbf{H}\mathbf{H}'du\right]^{-1}\int_0^1 \mathbf{H}(d\mathbf{B})'\right\}, \tag{42}$$

where $\mathbf{B}$ is a $2(p-r)-$dimensional complex Brownian motion, and $\mathbf{H}$ is some process derived from $\mathbf{B}$ depending on the model for the deterministic terms.

The Brownian motion $\mathbf{B}$ is approximated by a 400 - step random walk and the statistic is calculated 100.000 times or 500.000 times.

The approximation formulae used are as follows. Let $B=(B_R', B_I')'$ denote a $2(p-r)$-dimensional Brownian motion, and let $(\varepsilon_t)_{t\geq 0}$ be a sequence of $2(p-r)$-dimensional i.i.d. $N_{2(p-r)}(0,I)$ variables, then

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T}\left(\sum_{k=1}^{t}\varepsilon_k\right) \xrightarrow{w} \int_0^1 B(u)du,$$
$$\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{k=1}^{t-1}\varepsilon_k\right)\varepsilon_t' \xrightarrow{w} \int_0^1 B\,(dB)',$$
$$\frac{1}{T^2}\sum_{t=1}^{t}\left(\sum_{k=1}^{t}\varepsilon_k\right)\left(\sum_{k=1}^{t}\varepsilon_k\right)' \xrightarrow{w} \int_0^1 BB'du.$$

and

$$\mathbf{B}=\begin{pmatrix} B_R & -B_I \\ B_I & B_R \end{pmatrix}.$$

Table 1: Quantiles of the limit distribution for cointegration rank at seasonal frequency for model (10) with no determinstic terms, given by (42) with $\mathbf{H} = \mathbf{B}$. The number of iterations is 500.000, and the random walk has 400 steps.

p-r	0.01	0.05	0.10	0.50	0.75	0.80	0.85	0.90	0.95	0.975	0.99
1	0.0228	0.114	0.234	1.50	2.95	3.41	3.99	4.80	6.20	7.57	9.45
2	4.21	5.74	6.73	11.4	14.6	15.5	16.6	18.1	20.4	22.6	25.3
3	16.3	19.4	21.3	29.2	34.1	35.5	37.0	39.1	42.3	45.3	48.9
4	36.3	41.1	43.8	54.8	61.5	63.2	65.3	67.9	72.0	75.7	80.3
5	64.1	70.5	74.2	88.3	96.6	98.7	101	105	110	114	119
6	99.6	108	112	129	139	142	145	149	155	160	166
7	143	153	158	178	190	193	196	201	207	213	220
8	194	205	211	235	248	251	255	260	268	274	282
9	252	265	272	299	313	317	322	327	336	343	352
10	318	333	341	370	387	391	396	402	411	419	429
11	391	408	417	449	467	472	477	484	494	503	513
12	472	490	500	535	555	560	566	573	584	594	605

Table 2: Quantiles of the limit distribution for cointegration rank at seasonal frequency for model (11) with restricted seasonal dummies and constant, given by (42) with $\mathbf{H} = (\mathbf{B}', I)'$. The number of iterations is 100.000, and the random walk has 400 steps

p-r	0.01	0.05	0.10	0.50	0.75	0.80	0.85	0.90	0.95	0.975	0.99
1	0.457	1.06	1.63	5.91	9.72	10.8	12.2	13.9	16.9	19.8	23.5
2	10.4	13.6	15.6	25.3	32.4	34.5	37.1	40.8	46.9	53.0	61.4
3	30.5	36.1	39.7	55.5	66.9	70.3	74.4	80.2	89.6	99.1	112
4	60.0	68.6	73.7	96.5	113	117	123	131	144	157	174
5	99.3	111	118	148	169	175	183	193	210	226	247
6	148	163	172	210	237	244	253	266	286	305	331
7	206	224	235	283	315	324	335	350	374	397	428
8	273	296	309	366	404	415	428	446	473	501	537
9	352	377	393	459	504	516	532	552	584	615	657
10	438	468	486	563	615	629	646	669	705	741	786
11	534	569	590	678	735	751	771	797	838	877	927
12	639	679	703	802	867	885	907	936	982	1024	1078

Table 3: Quantiles of the limit distribution for cointegration rank at seasonal frequency for model (12) with restricted seasonal dummies and unrestricted constant, given by (42) with $\mathbf{H} = (\mathbf{B}' - \bar{\mathbf{B}}', \mathbf{I})'$. The number of iterations is 100.000, and the random walk has 400 steps

p-r	0.01	0.05	0.10	0.50	0.75	0.80	0.85	0.90	0.95	0.975	0.99
1	0.238	0.923	1.58	5.49	8.32	9.11	10.1	11.4	13.5	15.5	18.0
2	8.70	11.2	12.8	19.6	24.1	25.2	26.7	28.6	31.7	34.5	37.9
3	25.4	29.5	32.0	41.9	47.9	49.6	51.6	54.1	57.9	61.6	66.0
4	50.1	55.9	59.3	72.4	80.2	82.3	84.8	88.0	92.8	97.4	103
5	82.9	90.5	94.8	111	121	123	126	130	136	142	148
6	124	133	138	158	170	173	177	181	188	195	202
7	173	184	190	214	228	231	235	241	249	257	266
8	230	243	250	278	294	298	303	309	318	327	338
9	295	310	319	350	368	373	379	386	397	407	420
10	369	386	395	431	451	457	464	472	485	497	512
11	451	470	481	521	544	550	558	567	582	596	613
12	540	562	574	619	645	652	661	671	688	704	723

EUI WORKING PAPERS

EUI Working Papers are published and distributed by the
European University Institute, Florence

Copies can be obtained free of charge
– depending on the availability of stocks – from:

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

To

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI) – Italy
Telefax No: +39/55/4685 636
E-mail: publish@datacomm.iue.it

From

Name . . . .
Address. . . . .
. . . .
. . . .
. . . .
. . . .

❒ Please send me a complete list of EUI Working Papers
❒ Please send me a complete list of EUI book publications
❒ Please send me the EUI brochure Academic Year 1997/98

Please send me the following EUI Working Paper(s):

No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .

Date . . . .

Signature . . . .

Working Papers of the Department of Economics Published since 1994

ECO No. 96/1
Ana Rute CARDOSO
Earnings Inequality in Portugal: High and Rising?

ECO No. 96/2
Ana Rute CARDOSO
Workers or Employers: Who is Shaping Wage Inequality?

ECO No. 96/3
David F. HENDRY/Grayham E. MIZON
The Influence of A.W.H. Phillips on Econometrics

ECO No. 96/4
Andrzej BANIAK
The Multimarket Labour-Managed Firm and the Effects of Devaluation

ECO No. 96/5
Luca ANDERLINI/Hamid SABOURIAN
The Evolution of Algorithmic Learning: A Global Stability Result

ECO No. 96/6
James DOW
Arbitrage, Hedging, and Financial Innovation

ECO No. 96/7
Marion KOHLER
Coalitions in International Monetary Policy Games

ECO No. 96/8
John MICKLEWRIGHT/ Gyula NAGY
A Follow-Up Survey of Unemployment Insurance Exhausters in Hungary

ECO No. 96/9
Alastair McAULEY/John MICKLEWRIGHT/Aline COUDOUEL
Transfers and Exchange Between Households in Central Asia

ECO No. 96/10
Christian BELZIL/Xuelin ZHANG
Young Children and the Search Costs of Unemployed Females

ECO No. 96/11
Christian BELZIL
Contiguous Duration Dependence and Nonstationarity in Job Search: Some Reduced-Form Estimates

ECO No. 96/12
Ramon MARIMON
Learning from Learning in Economics

ECO No. 96/13
Luisa ZANFORLIN
Technological Diffusion, Learning and Economic Performance: An Empirical Investigation on an Extended Set of Countries

ECO No. 96/14
Humberto LÓPEZ/Eva ORTEGA/Angel UBIDE
Explaining the Dynamics of Spanish Unemployment

ECO No. 96/15
Spyros VASSILAKIS
Accelerating New Product Development by Overcoming Complexity Constraints

ECO No. 96/16
Andrew LEWIS
On Technological Differences in Oligopolistic Industries

ECO No. 96/17
Christian BELZIL
Employment Reallocation, Wages and the Allocation of Workers Between Expanding and Declining Firms

ECO No. 96/18
Christian BELZIL/Xuelin ZHANG
Unemployment, Search and the Gender Wage Gap: A Structural Model

ECO No. 96/19
Christian BELZIL
The Dynamics of Female Time Allocation upon a First Birth

ECO No. 96/20
Hans-Theo NORMANN
Endogenous Timing in a Duopoly Model with Incomplete Information

*out of print

ECO No. 96/21
Ramon MARIMON/Fabrizio ZILIBOTTI
'Actual' Versus 'Virtual' Employment in Europe: Is Spain Different?

ECO No. 96/22
Chiara MONFARDINI
Estimating Stochastic Volatility Models Through Indirect Inference

ECO No. 96/23
Luisa ZANFORLIN
Technological Diffusion, Learning and Growth: An Empirical Investigation of a Set of Developing Countries

ECO No. 96/24
Luisa ZANFORLIN
Technological Assimilation, Trade Patterns and Growth: An Empirical Investigation of a Set of Developing Countries

ECO No. 96/25
Giampiero M.GALLO/Massimiliano MARCELLINO
In Plato's Cave: Sharpening the Shadows of Monetary Announcements

ECO No. 96/26
Dimitrios SIDERIS
The Wage-Price Spiral in Greece: An Application of the LSE Methodology in Systems of Nonstationary Variables

ECO No. 96/27
Andrei SAVKOV
The Optimal Sequence of Privatization in Transitional Economies

ECO No. 96/28
Jacob LUNDQUIST/Dorte VERNER
Optimal Allocation of Foreign Debt Solved by a Multivariate GARCH Model Applied to Danish Data

ECO No. 96/29
Dorte VERNER
The Brazilian Growth Experience in the Light of Old and New Growth Theories

ECO No. 96/30
Steffen HÖRNIG/Andrea LOFARO/ Louis PHLIPS
How Much to Collude Without Being Detected

ECO No. 96/31
Angel J. UBIDE
The International Transmission of Shocks in an Imperfectly Competitive International Business Cycle Model

ECO No. 96/32
Humberto LOPEZ/Angel J. UBIDE
Demand, Supply, and Animal Spirits

ECO No. 96/33
Andrea LOFARO
On the Efficiency of Bertrand and Cournot Competition with Incomplete Information

ECO No. 96/34
Anindya BANERJEE/David F. HENDRY/Grayham E. MIZON
The Econometric Analysis of Economic Policy

ECO No. 96/35
Christian SCHLUTER
On the Non-Stationarity of German Income Mobility (and Some Observations on Poverty Dynamics)

ECO No. 96/36
Jian-Ming ZHOU
Proposals for Land Consolidation and Expansion in Japan

ECO No. 96/37
Susana GARCIA CERVERO
Skill Differentials in the Long and in the Short Run. A 4-Digit SIC Level U.S. Manufacturing Study

***

ECO No. 97/1
Jonathan SIMON
The Expected Value of Lotto when not all Numbers are Equal

ECO No. 97/2
Bernhard WINKLER
Of Sticks and Carrots: Incentives and the Maastricht Road to EMU

ECO No. 97/3
James DOW/Rohit RAHI
Informed Trading, Investment, and Welfare

*out of print

ECO No. 97/4
Sandrine LABORY
Signalling Aspects of Managers' Incentives

ECO No. 97/5
Humberto LÓPEZ/Eva ORTEGA/Angel UBIDE
Dating and Forecasting the Spanish Business Cycle

ECO No. 97/6
Yadira GONZÁLEZ de LARA
Changes in Information and Optimal Debt Contracts: The Sea Loan

ECO No. 97/7
Sandrine LABORY
Organisational Dimensions of Innovation

ECO No. 97/8
Sandrine LABORY
Firm Structure and Market Structure: A Case Study of the Car Industry

ECO No. 97/9
Elena BARDASI/Chiara MONFARDINI
The Choice of the Working Sector in Italy: A Trivariate Probit Analysis

ECO No. 97/10
Bernhard WINKLER
Coordinating European Monetary Union

ECO No. 97/11
Alessandra PELLONI/Robert WALDMANN
Stability Properties in a Growth Model

ECO No. 97/12
Alessandra PELLONI/Robert WALDMANN
Can Waste Improve Welfare?

ECO No. 97/13
Christian DUSTMANN/Arthur van SOEST
Public and Private Sector Wages of Male Workers in Germany

ECO No. 97/14
Søren JOHANSEN
Mathematical and Statistical Modelling of Cointegration

ECO No. 97/15
Tom ENGSTED/Søren JOHANSEN
Granger's Representation Theorem and Multicointegration

ECO No. 97/16
Søren JOHANSEN/
Ernst SCHAUMBURG
Likelihood Analysis of Seasonal Cointegration

*out of print